高等职业院校教学改革创新示范教材·网络开发系列

高级交换与路由技术

张国清　主　编

孙丽萍　杨　宇　崔升广　副主编

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

VISIT…

内 容 简 介

本书是一部关于网络设备配置与调试高级技术的项目训练教材。主要内容包括 PVLAN 技术与应用、多生成树协议 MSTP、虚拟路由器冗余协议 VRRP、动态主机配置协议 DHCP、AAA 机制与 RADIUS 应用、802.1X 安全访问控制、RIP 路由协议与高级配置、OSPF 路由协议与高级配置、路由选择控制与过滤、路由重发布、策略路由、BGP 路由协议与基本配置、GRE 隧道协议、IPSec 隧道协议 14 个训练项目。每个项目都以问题为导向，讲述相关知识及配置命令，通过应用案例演示设备配置与调试操作过程，学生完成相关操作训练后能够加深对相关知识的理解及提升相关操作技能。

本书既可以作为职业院校、应用型本科计算机网络技术专业理论与实训一体化教材，也可以作为社会职业培训教材，以及网络技术实训指导书。本书提供配套的电子课件、习题答案及两个补充参考的网络工程案例，请登录华信教育资源网（www.hxedu.com.cn）免费下载。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

高级交换与路由技术 / 张国清主编. —北京：电子工业出版社，2016.7
高等职业院校教学改革创新示范教材. 网络开发系列
ISBN 978-7-121-29250-7

Ⅰ. ①高… Ⅱ. ①张… Ⅲ. ①计算机网络－信息交换机－高等职业教育－教材②计算机网络－路由选择－高等职业教育－教材 Ⅳ. ①TN915.05

中国版本图书馆 CIP 数据核字（2016）第 149367 号

策划编辑：左 雅
责任编辑：左 雅
印 刷：
装 订：
出版发行：电子工业出版社
北京市海淀区万寿路 173 信箱 邮编 100036
开 本：787×1 092 1/16 印张：21 字数：604.8 千字
版 次：2016 年 7 月第 1 版
印 次：2016 年 7 月第 1 次印刷
印 数：3 000 册 定价：45.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）88254580 zuoya@phei.com.cn。

前　言

计算机网络技术是发展最快、应用最广泛的一项技术，极大地改变着人们的生产、生活方式。高级交换及路由技术是计算机网络技术中的核心内容，涉及内容深、范围广，既是重点又是难点。为了使学员能够真正掌握这一门技术，我们在本书中采用了面向问题的工作导向、面向过程的工作流程、面向职业的技能考核。通过图、表、文字，将高深、复杂、难懂的技术问题简单化；通过教师讲述、操作演示、技能训练等环节，将理论与实践相结合，循序渐进地让学生掌握操作技能。

电子工业出版社 2015 年出版发行的《网络设备配置与调试项目实训》（第 3 版）是在广泛听取读者意见和建议，充分总结第 1 版及第 2 版教材的基础上开发的一部适合职业教育类的网络设备配置项目训练教材。该教材是针对初次接触网络设备的学生而开发的，教材内容主要涉及交换和路由的基本知识及基本配置技能，可以作为本教材的前期专业基础课程教材。

《高级交换与路由技术》是在前期完成《计算机网络原理》及《网络互联技术》等专业基础课程学习后，提升网络设备配置与调试高级技术能力的项目训练教材。教材内容包括 PVLAN 技术与应用、多生成树协议 MSTP、虚拟路由器冗余协议 VRRP、动态主机配置协议 DHCP、AAA 机制与 RADIUS 应用、802.1X 安全访问控制、RIP 路由协议与高级配置、OSPF 路由协议与高级配置、路由选择控制与过滤、路由重发布、策略路由、BGP 路由协议与基本配置、GRE 隧道协议、IPSec 隧道协议 14 个训练项目。每个项目都以问题为导向，讲述相关知识及配置命令，通过应用案例演示设备配置与调试操作过程，学生完成相关操作训练后能够深入理解相关知识及提升相关操作技能水平。

在本书编写过程中，编者总结了多年计算机网络工程实践及高职教学经验，对网络设备配置与调试涉及的知识和技能进行了科学、合理的划分，优化《网络设备配置与调试项目实训》及《高级交换与路由技术》两部教材内容，根据实际工作所需知识和技能抽象出若干教学项目，最终形成为职业院校学生量身定做的计算机网络技术专业高级项目课程教材。

本书以知识“必需、够用”为原则，从职业岗位分析入手展开教学内容，强化学生的技能训练，在训练过程中巩固所学知识。全书以问题为导向，引出为完成工作任务需要学习的相关知识及掌握的相关技能，以问题解决方案的完成效果检验知识学习及技能掌握情况。

本书既可以作为职业院校、应用型本科计算机网络技术专业理论与实训一体化教材，也可以作为社会职业培训教材，还可以作为网络技术实训指导书。本书实训环境以锐捷公司的网络设备搭建网络环境为主要训练平台。在实际应用过程中，教师可以根据本校的网络实训环境做适当调整。

本书在编写过程中得到了星网锐捷网络有限公司张选波高级工程师的大力支持，在此表示深深的谢意。本书由张国清教授任主编，由孙丽萍、杨宇、崔升广任副主编，其中，项目 1、项目 2 由孙丽萍编写，项目 3 由杨宇编写，项目 4 由崔升广编写，项目 5 由刘莉莉编写，项目 6～项目 14 由张国清编写。全书由张国清统稿。

由于编者水平有限，书中难免有不妥和错误之处，恳请广大读者批评指正，联系邮箱地址为 Zgq8163@163.com。

编　者

目　　录

项目 1　PVLAN 技术与应用 …… 1
1.1　问题提出 …… 1
1.2　相关知识 …… 1
1.3　扩展知识 …… 7
1.4　演示实例 …… 8
1.5　训练任务 …… 11
练习题 …… 11
项目 2　多生成树协议 MSTP …… 13
2.1　问题提出 …… 13
2.2　相关知识 …… 13
2.3　扩展知识 …… 20
2.4　演示实例 …… 21
2.5　训练任务 …… 30
练习题 …… 31
项目 3　虚拟路由器冗余协议 VRRP …… 32
3.1　问题提出 …… 32
3.2　相关知识 …… 33
3.3　扩展知识 …… 41
3.4　演示实例 …… 44
3.5　训练任务 …… 47
练习题 …… 48
项目 4　动态主机配置协议 DHCP …… 49
4.1　问题提出 …… 49
4.2　相关知识 …… 49
4.3　扩展知识 …… 58
4.4　演示实例 …… 62
4.5　训练任务 …… 64
练习题 …… 65
项目 5　AAA 机制与 RADIUS 应用 …… 67
5.1　问题提出 …… 67
5.2　相关知识 …… 67
5.2.1　AAA 概述 …… 67
5.2.2　配置 AAA 的验证功能 …… 68
5.2.3　RADIUS 服务 …… 75
5.3　扩展知识 …… 82
5.3.1　配置 AAA 的授权功能 …… 82
5.3.2　配置 AAA 的记账功能 …… 82

5.4 演示实例 …… 83
5.5 训练任务 …… 84
练习题 …… 85
项目 6 802.1X 安全访问控制 …… 86
6.1 问题提出 …… 86
6.2 相关知识 …… 86
6.3 扩展知识 …… 93
6.4 演示实例 …… 96
6.5 训练任务 …… 99
练习题 …… 100
项目 7 RIP 路由协议与高级配置 …… 101
7.1 问题提出 …… 101
7.2 相关知识 …… 101
7.3 知识扩展 …… 111
7.4 演示实例 …… 112
7.5 训练任务 …… 114
练习题 …… 114
项目 8 OSPF 路由协议与高级配置 …… 116
8.1 问题提出 …… 116
8.2 相关知识 …… 116
8.2.1 回顾 OSPF 协议的基本特性及应用配置 …… 116
8.2.2 OSPF 协议特性及高级配置 …… 130
8.3 扩展知识 …… 135
8.4 演示实例 …… 147
8.5 训练任务 …… 148
练习题 …… 150
项目 9 路由选择控制与过滤 …… 152
9.1 问题提出 …… 152
9.2 相关知识 …… 152
9.2.1 被动接口 …… 152
9.2.2 分发列表 …… 158
9.3 扩展知识 …… 166
9.4 演示实例 …… 170
9.5 训练任务 …… 174
练习题 …… 177
项目 10 路由重发布 …… 179
10.1 问题提出 …… 179
10.2 相关知识 …… 179
10.3 扩展知识 …… 186
10.4 演示实例 …… 192
10.5 训练任务 …… 198

练习题……203
项目 11　策略路由……205
11.1　问题提出……205
11.2　相关知识……205
11.3　扩展知识……212
11.4　演示实例……214
11.5　训练任务……219
练习题……220
项目 12　BGP 路由协议与基本配置……222
12.1　问题提出……222
12.2　相关知识……222
练习题……283
项目 13　GRE 隧道协议……286
13.1　问题提出……286
13.2　相关知识……286
13.3　扩展知识……291
13.4　演示实例……291
13.5　训练任务……295
练习题……295
项目 14　IPSec 隧道协议……296
14.1　问题提出……296
14.2　相关知识……296
14.3　扩展知识……317
14.4　演示实例……323
14.5　训练任务……326
练习题……326
参考文献……327

项目 1　PVLAN 技术与应用

通过前面的学习，我们知道划分 VLAN 可以隔离广播域，一台交换机内最多可以划分 4094 个 VLAN，如果需要更多的广播域，就会受到限制。另一方面，每一个 VLAN 需要一个子网地址，导致地址浪费。如何来解决这个问题呢？使用 PVLAN 技术可以解决。

通过本章学习能够对交换机 PVLAN 有所了解，掌握交换机 PVLAN 配置技能。

知识点、技能点

1. 了解 PVLAN 概念;
2. 掌握 PVLAN 特性及其配置技能;
3. 掌握小型网络一般调试技能及故障排除方法。

1.1　问题提出

为了在二层网络上隔离用户或者隔离广播，我们可以将一组设备加入到一个 VLAN 中，但是 VLAN 的最大个数只有 4094，所以当我们需要隔离大量的广播域时将可能受到 VLAN 个数的限制。通常，在 Internet 服务提供商（ISP）网络中，为了隔离不同客户之间的通信而将一个客户作为一个 VLAN，但是如果客户的数量增大到 VLAN 的最大个数（4094）时，服务提供商提供的服务将会受到限制。而且，对于一个客户作为一个 VLAN 的解决方案，服务提供商需要为一个客户分配一个子网地址，这种情况导致了 IP 地址的浪费。对于这些问题我们可以使用 Private VLAN 技术来解决。

1.2　相关知识

1. PVLAN 概述

PVLAN 是 Private VLAN 的缩写，即私有 VLAN。PVLAN 是能够为相同 VLAN 内不同端口提供隔离的 VLAN。通过 PVLAN 技术可以隔离相同 VLAN 中的网络设备之间的流量，并且位于相同子网的所有设备都只能与网关或其他网络进行通信，实现网络内部的隔离。

PVLAN 可以将一个 VLAN 的二层广播域划分成多个子域，每个子域都由一对 VLAN 组成：Primary VLAN（主 VLAN）和 Secondary VLAN（辅助 VLAN）。在整个 PVLAN 域中，只有一个主 VLAN，每个子域有不同的辅助 VLAN，通过辅助 VLAN 实现二层网络的隔离。

（1）主 VLAN：主 VLAN 是 PVLAN 的高级 VLAN，每个 PVLAN 中只有一个主 VLAN。

（2）辅助 VLAN：辅助 VLAN 是 PVLAN 中的子 VLAN，并且映射到一个主 VLAN。每台接入设备都连接到辅助 VLAN 中。

辅助 VLAN 包含以下两种类型。

（1）隔离 VLAN（Isolated VLAN）：同一个隔离 VLAN 中的端口互相不能进行二层通信，一个私有 VLAN 域中只有一个隔离 VLAN。

（2）团体 VLAN（Community VLAN）：同一个团体 VLAN 中的端口可以进行二层通信，但是不能与其他团体 VLAN 中的端口进行二层通信，一个私有 VLAN 域中可以有多个团体 VLAN。

PVLAN 中的端口有如下几种类型。

（1）混杂端口（Promiscuous Port）：混杂端口为主 VLAN 中的端口，可以与任意端口通信，包括同一个 PVLAN 中的隔离端口和团体端口。

（2）隔离端口（Isolated Port）：隔离端口为隔离 VLAN 中的端口，隔离端口只能与混杂端口进行通信。

（3）团体端口（Community Port）：团体端口为团体 VLAN 中的端口，同一个团体 VLAN 中的团体端口之间可以互相通信，并且团体端口可以与混杂端口通信，但是不能与其他团体 VLAN 的端口进行通信。

PVLAN 中主 VLAN 与辅助 VLAN 间关系、各类辅助 VLAN 特性、各类端口特性如图 1.1 所示。

在 PVLAN 网络中，通过 PVLAN 技术针对不同的部门配置不同的 VLAN 类型，从而实现部门间的隔离或部门内部主机的隔离。

在如图 1.2 所示的网络中，在主 VLAN 中有三个部门，其中行政部属于团体 VLAN10，商务部属于团体 VLAN20，财务部属于隔离 VLAN30。由于三个部门属于同一个主 VLAN，因此三个部门中主机的 IP 都属于同一个子网，而行政部里的各设备能互相通信，商务部里的各设备也能互相通信，而财务部门的设备由于属于隔离 VLAN，因此互相之间不能互相通信。但是三个部门的各设备都能与主 VLAN 中的混杂端口通信从而实现对其他网络的访问。

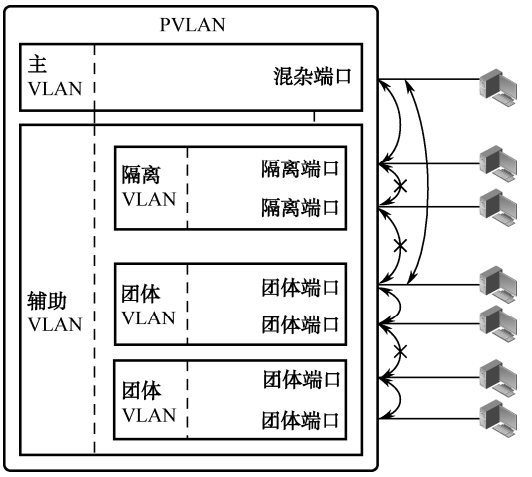


图 1.1　PVLAN 功能特性及应用

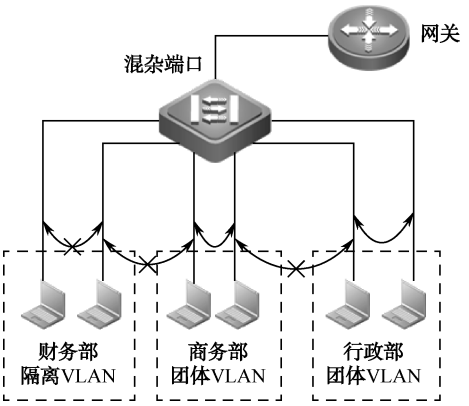


图 1.2　PVLAN 应用实例

2. 配置 PVLAN

默认情况下，交换机没有 Private VLAN 的配置。若要配置 Private VLAN 则需要以下几个步骤。

第 1 步：创建主 VLAN 与辅助 VLAN。

第 2 步：关联辅助 VLAN 到主 VLAN。

第 3 步：将辅助 VLAN 映射到主 VLAN 的三层端口。

第 4 步：配置主机端口（Host Port）。

第 5 步：配置混杂端口。

（1）创建主 VLAN 与辅助 VLAN。

默认情况下，交换机只有一个 VLAN1，交换机的全部端口都属于 VLAN1 内的端口。为了实现 PVLAN 的功能，必须先创建 VLAN，然后设置 PVLAN 类型是 Primary VLAN、Community VLAN 及 Isolated VLAN。

设置 PVLAN 类型的命令格式如下：

其中，*community* 为配置 Community VLAN；*isolated* 为配置 Isolated VLAN；*primary* 为配置 Primary VLAN。

提示：*使用 no private-VLAN {community | isolated | primary}命令取消相应的私有 VLAN 配置，如使用 ruijie(config-VLAN)#no private-VLAN community 命令取消团体 VLAN 配置。*

在 802.1Q VLAN 中，有成员端口的 VLAN 不能配置为私有 VLAN。VLAN1 不能配置为私有 VLAN。

例如：设置 VLAN99 为 Primary VLAN、VLAN100 为 Community VLAN、VLAN101 为 Isolated VLAN。

```
ruijie(config)#VLAN 99
ruijie(config-VLAN)#private-VLAN primary
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 100
ruijie(config-VLAN)#private-VLAN community
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 101
ruijie(config-VLAN)#private-VLAN isolated
ruijie(config-VLAN)#exit
ruijie(config)#
```

例如：取消之前的 VLAN101 为 Isolated VLAN 的设置。

```
ruijie(config)#VLAN 101
ruijie(config-VLAN)#no private-VLAN isolated
ruijie(config-VLAN)#exit
```

使用 show VLAN private-VLAN [community | primary | isolated]命令查看当前 PVLAN 配置情况如下：

```
ruijie#show VLAN private-VLAN
VLAN   Type        Status     Routed      Interface    Associated VLANs
----    --------    --------   ------      ---------     ----------------
99      primary     inactive   Disabled
100     community   inactive   Disabled                  No Association
101     isolated    inactive   Disabled                  No Association
```

从上述显示结果可以看出，设置了 PVLAN 类型，但没有建立关联关系。

（2）关联 Secondary（辅助）VLAN 到 Primary（主）VLAN。

设置 PVLAN 类型后，还需要进入 Primary VLAN 配置模式下，将辅助 VLAN 关联到主 VLAN。关联辅助 VLAN 到主 VLAN 的命令格式如下：

设置主VLAN关联辅助VLAN　　*添加辅助VLAN列表*

ruijie(config-VLAN)#private-VLAN association {*svlist* **| add** *svlist* **| remove** *svlist***}**

交换机名称为ruijie
处于VLAN配置模式　　*关联辅助VLAN列表*　　*删除辅助VLAN列表*

其中，*svlist* 为辅助 PVLAN 列表。

提示：*使用 no private-VLAN association 命令清除与所有辅助 VLAN 的关联。*

例如：将 Secondary VLAN 中的 Community VLAN 类型的 VLAN100、Isolated VLAN 类型的 VLAN101 关联到作为 Primary VLAN 的 VLAN99 中。

```
ruijie(config)#VLAN 99
ruijie(config-VLAN)#private-VLAN association 100,101
ruijie(config-VLAN)#exit
ruijie(config)#
```

使用 show VLAN private-VLAN [community | primary | isolated]命令查看当前 PVLAN 配置情况如下：

```
ruijie#show VLAN private-VLAN
VLAN    Type        Status      Routed     Ports    Associated VLANs
-----   ----------  --------    --------   ------   ------------------
99      primary     active            Disabled     100-101
100     community   active            Disabled     99
101     isolated    active            Disabled     99
```

从上述显示结果可以看出，主 VLAN99 与辅助 VLAN100、VLAN101 建立了关联关系。

例如：清除主 VLAN 的 VLAN99 与所有辅助 VLAN 的关联。

```
ruijie(config)#VLAN 99
ruijie(config-VLAN)#no private-VLAN association
ruijie(config-VLAN)#exit
ruijie(config)#
```

使用 show VLAN private-VLAN [community | primary | isolated]命令查看当前 PVLAN 配置情况如下：

```
ruijie#show VLAN private-VLAN
VLAN   Type        Status     Routed       Interface      Associated VLANs
----    --------   --------   ------       ---------      ----------------
99      primary    inactive   Disabled
100     community  inactive   Disabled                    No Association
101     isolated   inactive   Disabled                    No   Association
```

从上述显示结果可以看出，主 VLAN 与辅助 VLAN 之间的关联关系都已取消了。

（3）将辅助 VLAN 映射到主 VLAN 的 SVI 端口。

为了使 PVLAN 中的流量可以进行三层路由，需要为主 VLAN 配置 SVI，并且将辅助 VLAN 映射到主 VLAN 的 SVI。为了将辅助 VLAN 映射到主 VLAN 的三层端口，必须先进入主 VLAN 端口，然后在主 VLAN 端口下使用如下命令完成该映射任务：

将辅助VLAN映射到主VLAN的SVI端口　*添加辅助VLAN列表*

ruijie(config-If)#private-VLAN mapping {*svlist* **| add** *svlist* **| remove** *svlist***}**

交换机名称为ruijie
处于端口配置模式　*关联辅助VLAN列表*　*删除辅助VLAN列表*

其中，*svlist* 为辅助 PVLAN 列表。

提示：*使用 no private-VLAN mapping 命令清除与所有辅助 VLAN 的映射关系。*

例如：VLAN99 为主 VLAN 且 SVI 的 IP 地址为 192.168.1.1，VLAN100 为团体 VLAN，VLAN101 为隔离 VLAN。现将辅助 VLAN（VLAN100、VLAN101）映射到主 VLAN 的三层端口上。

```
ruijie(config)#interface VLAN 99
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
ruijie(config-if)#private-VLAN mapping 100-101
ruijie(config-if)#exit
ruijie(config)#
```

例如：取消已建立的到主 VLAN 的 SVI 端口的映射关系。

```
ruijie(config)#interface VLAN 99
ruijie(config-if)#no private-VLAN mapping
ruijie(config-if)#exit
```

（4）配置主机端口。

在 PVLAN 中交换机端口分为两大类：混杂端口和主机端口。其中，主机端口又分为隔离端口和团体端口。它们主要作为主机设备的接入端口使用。配置 Host Port 操作时需要 2 个配置步骤：第 1 步，设置端口类型；第 2 步，将主机端口关联到 Primary VLAN 和 Secondary VLAN 上。具体配置 Host Port 的步骤如下。

第 1 步：设置 PVLAN 端口类型为 Host Port。

设置 PVLAN 端口类型为 Host Port 的命令如下：

设置PVLAN的端口类型　*混杂端口类型*

ruijie(config-If)#switchport mode private-VLAN{*host*|*promiscuous***}**

交换机名称为ruijie
处于端口配置模式　*主机端口类型*

其中，*host* 为私有 VLAN 的主机端口类型；*promiscuous* 为私有 VLAN 的混杂端口类型。

提示：*使用 no switchport mode 命令删除端口私有 VLAN 配置。*

例如：设置端口 FastEthernet 0/10 为 Host Port 类型。

```
ruijie(config)#interface FastEthernet0/10
ruijie(config-if)#switchport mode private-VLAN host
```

例如：删除端口 FastEthernet 0/10 的私有 VLAN 配置。

```
ruijie(config)#interface FastEthernet 0/10
ruijie(config-if)#no switchport mode
```

第 2 步：将主机端口关联到 Primary VLAN 和 Secondary VLAN。

设置完 PVLAN 端口类型后，需要将该端口关联到 Primary VLAN 和 Secondary VLAN 上，完成该关联任务的命令如下：

设置PVLAN的主机端口所关联的主VLAN及辅助VLAN　　辅助VLAN编号

ruijie(config-If)#switchport private-VLAN host-association *p_vid s_vid*

交换机名称为ruijie
处于端口配置模式　　主VLAN编号

其中，*p_vid* 为已经创建的 Primary VID；*s_vid* 为已经创建的 Secondary VID。

提示：*使用 no switchport private-VLAN host-association 命令从 PVLAN 中删除主机端口。*

例如：将主机端口 FastEthernet 0/10 关联到作为 Primary VLAN 的 VLAN99 和作为 Community VLAN 的 VLAN100。

```
ruijie(config)#interface FastEthernet 0/10
ruijie(config-if)#switchport mode private-VLAN host
ruijie(config-if)#switchport private-VLAN host-association 99 100
ruijie(config-if)#exit
ruijie(config)#
```

例如：将主机端口 FastEthernet 0/10 从 PVLAN 中删除。

```
ruijie(config)#interface FastEthernet 0/10
ruijie(config-if)#no switchport private-VLAN host-association
ruijie(config-if)#exit
ruijie(config)#
```

（5）配置混杂端口（Promiscuous Port）。

配置 Promiscuous Port 同配置主机端口（Host Port）一样，也需要 2 个操作步骤：第 1 步，设置端口类型；第 2 步，将混杂端口关联到 Primary VLAN 和 Secondary VLAN 上。具体配置 Promiscuous Port 的步骤如下。

第 1 步：设置 PVLAN 端口类型为 Promiscuous Port。

设置 PVLAN 端口类型为 Promiscuous Port 的命令格式如下：

设置PVLAN的端口类型　　混杂端口类型

ruijie(config-If)#switchport mode private-VLAN{*host*|*promiscuous***}**

交换机名称为ruijie
处于端口配置模式　　主机端口类型

其中，*host* 为私有 VLAN 的主机端口类型；*Promiscuous* 为私有 VLAN 的混杂端口类型。

提示：*使用 no switchport mode 命令删除端口私有 VLAN 配置。*

例如：配置 FastEthernet 0/1 为 Promiscuous Port。

```
ruijie(config)#interface FastEthernet 0/1
ruijie(config-if)#switchport mode private-VLAN promiscuous
ruijie(config-if)#end
ruijie#
```

第 2 步：将混杂端口关联到 Primary VLAN 和 Secondary VLAN 上。

将 Promiscuous Port 关联到 Primary VLAN 和 Secondary VLAN 上的命令格式如下：

设置PVLAN的混杂端口所关联的主VLAN及辅助VLAN　　辅助VLAN编号

ruijie(config-If)#switchport private-VLAN mapping *p_vid***{***svlist* **| add** *svlist* **| remove** *svlist*

交换机名称为ruijie
处于端口配置模式　　主VLAN编号

其中，*p_vid* 为已经创建的 Primary VID；*svlist* 为已经创建的 Secondary VLAN list。

提示：*使用 no switchport private-VLAN mapping 命令从 PVLAN 中删除混杂端口。*

例如：配置 FastEthernet 0/1 为 Promiscuous Port，并且关联到 Primary VLAN（VLAN99）、Community VLAN（VLAN100）和 Isolated VLAN（VLAN101）。

```
ruijie(config)#interface FastEthernet 0/1
ruijie(config-if)#switchport mode private-VLAN    promiscuous
ruijie(config-if)#switchport private-VLAN mapping 99 add 100-101
ruijie(config-if)#exit
ruijie(config)#
```

例如：将混杂端口 FastEthernet 0/1 从 PVLAN 中删除。

```
ruijie(config)#interface FastEthernet 0/1
ruijie(config-if)#no switchport private-VLAN mapping
ruijie(config-if)#exit
ruijie(config)#
```

（6）显示 PVLAN 配置信息。

显示 Private VLAN 的配置信息的命令格式如下：

显示PVLAN　　*辅助VLAN编号*

ruijie#show VLAN private-VLAN *[community | primary | isolated]*

交换机名称为ruijie
处于特权模式　　*团体VLAN*

其中，*primary* 为显示 Primary VLAN 信息；*community* 为显示 Community VLAN 信息；*Isolated* 为显示 Isolated VLAN 信息。

1.3　扩展知识

1. 保护端口（Protected Port）概述

在某些应用环境下，要求一台交换机上的有些端口之间不能互相通信。在这种环境下，通过设置保护口（Protected Port）实现这些端口之间拒绝单播帧、多播帧、广播帧的链路层通信，这些端口之间只有通过三层设备进行通信。保护端口之间无法通信，保护端口与非保护端口之间可以通信，如图 1.3 所示。

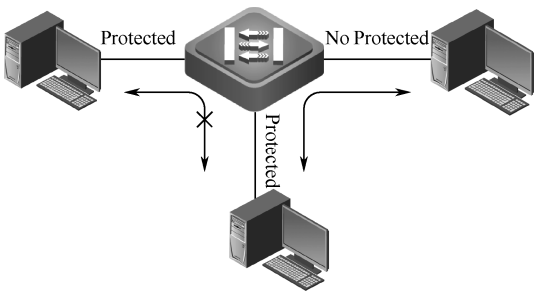


图 1.3　Protected Port 功能

只能实现同一台交换机的相同 VLAN 内的端口链路层通信隔离，不能隔离位于不同交换机或不同

VLAN 的保护端口之间的通信。在同一台交换机的相同 VLAN 内的保护端口之间可以通过路由端口通信。

前面介绍的 PVLAN 隔离 VLAN 内的端口，不但 VLAN 内的端口之间不能通信，不同 VLAN 之间的端口也不能通信，只能与混杂端口通信。

2. 配置保护端口

为了配置交换机保护端口，需要先进入端口模式，然后在端口模式下配置保护端口。配置保护端口的命令格式如下：

设置保护端口

ruijie(config-If)#switchport protected

交换机名称为ruijie
处于端口配置模式

提示：*使用 no switchport protected 取消保护端口设置，如 ruijie(config-if)no switchport protected。*

例如：将交换机端口 Fa0/1-12 设置为保护端口。

```
ruijie(config)#interface range fa 0/1-12
ruijie(config-range-if)#switchport protected
ruijie(config-range-if)#
```

1.4 演示实例

1. 背景描述

在如图 1.4 所示的网络中，在交换机上创建 VLAN10，并设置为 Primary VLAN，设置其 SVI 端口 IP 地址为 192.168.10.1/24；创建 VLAN20，并设置为 Isolated VLAN；创建 VLAN30，并设置为 Community VLAN；创建 VLAN40，并设置 VLAN40 的 SVI 端口 IP 地址为 192.168.40.1/24。实现团体 VLAN30、隔离 VLAN20 中的计算机能够与混杂端口 Fa0/1 端的 Server 通信；也能够与另一个网络 VLAN40 中的 PC5 通信。隔离 VLAN20 内的 PC1 与 PC2 之间也不能相互通信，也不能与团体 VLAN30 内的计算机通信。团体 VLAN30 内部的 PC3 与 PC4 之间能够相互通信，不能与隔离 VLAN20 内的 PC1 或 PC2 通信。

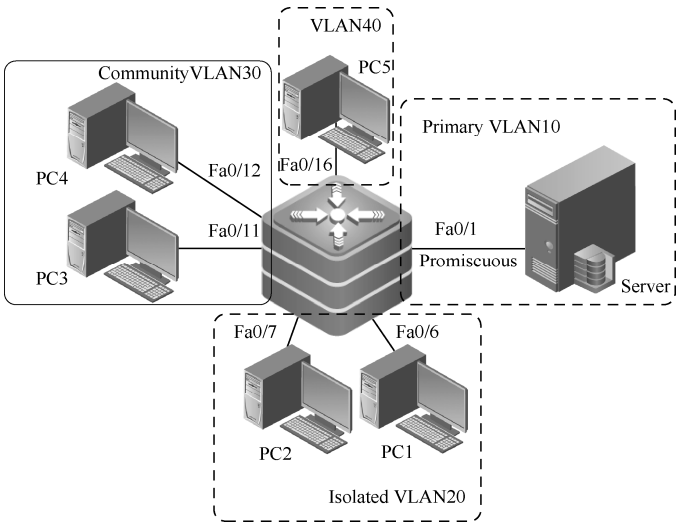


图 1.4　PVLAN 环境 SVI 应用

2. 操作步骤

第 1 步：创建主 VLAN 及辅助 VLAN。

```
ruijie(config)#VLAN 10
ruijie(config-VLAN)#private-VLAN primary
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 20
ruijie(config-VLAN)#private-VLAN isolated
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 30
ruijie(config-VLAN)#private-VLAN community
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 40
ruijie(config-VLAN)#end
ruijie#show VLAN private-VLAN
VLAN    Type          Status      Routed    Ports          Associated VLANs
-----   -------       --------    --------  -------------- ------------------
10      primary       inactive    Disabled
20      isolated      inactive    Disabled                 No Association
30      community     inactive    Disabled                 No Association
```

第 2 步：将辅助 VLAN 与主 VLAN 关联。

```
ruijie(config)#VLAN 10
ruijie(config-VLAN)#private-VLAN association 20,30
ruijie(config-VLAN)#end
ruijie#show VLAN private-VLAN
VLAN      Type          Status        Routed    Ports          Associated VLANs
-----     ----------    --------      --------  -------------- ------------------
10        primary       active        Disabled                 20,30
20        isolated      active        Disabled                 10
30        community     active        Disabled                 10
```

第 3 步：将辅助 VLAN 映射到 SVI 端口。

```
ruijie(config)#interface VLAN 10
ruijie(config-if)#ip address 192.168.10.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#private-VLAN mapping 20,30
ruijie(config-if)#end
ruijie#show VLAN private-VLAN
VLAN      Type          Status      Routed      Ports            Associated VLANs
-----     ----------    --------    --------    ---------------- ------------------
10        primary       active      Enabled                      20,30
20        isolated      active      Enabled                      10
30        community     active      Enabled                      10
```

第 4 步：配置隔离及团体端口。

```
ruijie(config)#interface range fa 0/6-7
```

```
ruijie(config-range-if)#switchport mode private-VLAN host
ruijie(config-range-if)#switchport private-VLAN host-association 10 20
ruijie(config-range-if)#exit
ruijie(config)#interface range fa 0/11-12
ruijie(config-range-if)#switchport mode private-VLAN host
ruijie(config-range-if)#switchport private-VLAN host-association 10 30
ruijie(config-range-if)#exit
ruijie#show VLAN private-VLAN
VLAN    Type        Status    Routed    Ports              Associated VLANs
-----   ----------  --------  --------  ----------------   ------------------
10      primary     active    Enabled                      20,30
20      isolated    active    Enabled   Fa0/6, Fa0/7       10
30      community   active    Enabled   Fa0/11, Fa0/12     10
```

第 5 步：配置混杂端口。

```
ruijie(config)#interface fa 0/1
ruijie(config-if)#switchport mode private-VLAN    promiscuous
ruijie(config-if)#switchport private-VLAN mapping 10 add 20,30
ruijie(config-if)#end
ruijie#show VLAN private-VLAN
VLAN    Type        Status    Routed    Ports              Associated VLANs
-----   ----------  --------  --------  ----------------   ------------------
10      primary     active    Enabled   Fa0/1              20,30
20      isolated    active    Enabled   Fa0/6, Fa0/7       10
30      community   active    Enabled   Fa0/11, Fa0/12     10
```

第 6 步：创建 VLAN40 的 SVI 端口。

```
ruijie(config)#interface fa 0/16
ruijie(config-if)#switchport access VLAN 40
ruijie(config-if)#exit
ruijie(config)#interface VLAN 40
ruijie(config-if)#ip address 192.168.40.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#end
ruijie#show ip    route
Codes:    C - connected, S - static, R - RIP, B - BGP
          O - OSPF, IA - OSPF inter area
          N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
          E1 - OSPF external type 1, E2 - OSPF external type 2
          i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
          ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       192.168.10.0/24 is directly connected, VLAN 10
C       192.168.10.1/32 is local host.
C       192.168.40.0/24 is directly connected, VLAN 40
```

```
C       192.168.40.1/32 is local host.
```

第 7 步：测试。

Server 的 IP 地址为 192.168.10.10/24；PC1 的 IP 地址为 192.168.10.11/24；PC2 的 IP 地址为 192.168.10.12/24；PC3 的 IP 地址为 192.168.10.13/24；PC4 的 IP 地址为 192.168.10.14/24，它们的网关地址为 192.168.10.1。PC5 的 IP 地址为 192.168.40.10/24，网关 IP 地址为 192.168.40.1。

PC1（PC2）能够与 Server、PC5 通信，但 PC1 与 PC2 之间不能通信；PC3（PC4）能够与 Server、PC5 通信，PC3 与 PC4 之间也能通信。

1.5　训练任务

1. 背景描述

在如图 1.5 所示的网络中，VLAN10 为 Primary VLAN，Secondary VLAN20 类型为 Community，Secondary VLAN30 类型为 Isolated。端口 Fa0/1 和 Fa0/2 为 VLAN20 的端口，端口 Fa0/11 和端口 Fa0/12 为 VLAN30 的端口。端口 Fa0/24 为 Private VLAN 的混杂端口，与 VLAN20 和 VLAN30 关联。

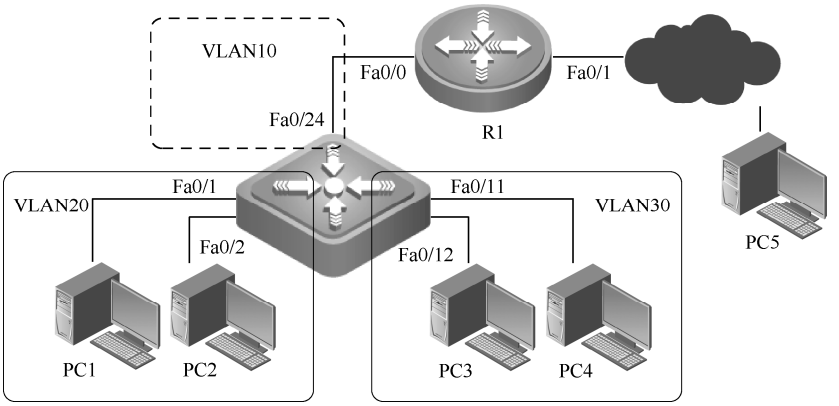


图 1.5　PVLAN 环境连接路由器应用

2. 操作提示

在此项目中，主 VLAN 只需设置 FastEthernet 0/24 为混杂端口，无须建立 VLAN10 的 SVI 端口及配置相应的 IP 地址。路由器 R1 的 FastEthernet 0/0 端口 IP 地址作为 VLAN20、VLAN30 中的主机的默认网关 IP 地址。其他 PVLAN 参数正常配置即可。

```
switch(config)#interface fa 0/24
switch(config-if)#switchport mode private-VLAN    promiscuous
switch(config-if)#switchport private-VLAN mapping 10 add 20,30
```

练习题

1. 选择题

（1）下面关于 PVLAN 描述正确的是（　　）。

A．主 VLAN 是 PVLAN 的高级 VLAN，每个 PVLAN 中只有一个主 VLAN。辅助 VLAN 是 PVLAN 中的子 VLAN，并且映射到一个主 VLAN

B．PVLAN 中的端口有三种类型：混杂端口、隔离端口和团体端口，混杂端口为主 VLAN 中的端口，隔离端口为隔离 VLAN 中的端口，团体端口为团体 VLAN 中的端口

C．辅助 VLAN 有两种类型：隔离 VLAN 和团体 VLAN，同一个团体 VLAN 中的端口可以进行二层通信，但是不能与其他团体 VLAN 中的端口进行二层通信

D．混杂端口可以与任意端口通信，包括同一个 PVLAN 中的隔离端口和团体端口，隔离端口只能与团体端口进行通信

（2）一个 Private VLAN 处于 Active 状态必须满足的条件为（　　）。

A．具有主 VLAN

B．具有辅助 VLAN

C．辅助 VLAN 和主 VLAN 可以进行关联

D．主 VLAN 必须设置 IP 地址

E．主 VLAN 内有混杂端口

2. 简答题

（1）请描述 PVLAN 的功能及应用场合。

（2）请描述 PVLAN 的配置命令及步骤。

（3）请说明 PVLAN 中各子 VLAN 中 IP 地址是如何分配的。

（4）请举例说明 PVLAN 是如何实现与其他网段的路由功能的。

项目 2　多生成树协议 MSTP

STP 及 RSTP 很好地解决了网络链路环路问题，使网络链路实现冗余备份，但是并没有考虑 VLAN 及负载均衡问题。要解决这个问题，需要采用多生成树协议（MSTP）。

通过本章学习能够对交换机 MSTP 定义有所了解，掌握交换机 MSTP 特性及其配置技能。

知识点、技能点

1. 了解 MSTP 概念;
2. 掌握 MSTP 特性及配置技能;
3. 掌握小型网络一般调试技能及故障排除方法。

2.1　问题提出

公司网络通过采用 RSTP 技术后，交换机端口状态切换时间明显缩短了。公司网络经过认真维护一直稳定运行。旧的问题解决了，新的问题又出现了。经过一段时间发现公司的网络流量都集中在两台核心交换机中的一台根交换机上，所有链路都以它为中心向外拓展，另一台非根交换机工作量较少，负载明显不平衡。必须采取措施，解决此问题。

2.2　相关知识

1. 多生成树协议（MSTP）概述

MSTP（Multiple Spanning Tree Protocol，多生成树协议）是在传统的 STP、RSTP 的基础上发展而来的新的生成树协议。MSTP 遵循的标准是 IEEE 802.1S。由于传统的生成树协议与 VLAN 没有任何联系，因此在特定网络拓扑下就会产生以下问题。

如图 2.1 所示，交换机 SW1、SW3 在 VLAN10 内，交换机 SW2、SW4 在 VLAN20 内，然后连成环路。

假设 SW2 的优先级为 4096、SW4 的优先级为 8192、SW1 及 SW3 的优先级为默认值 32768，交换机 SW2 成为根交换机，交换机 SW1 和 SW3 间的链路为 Discarding，如图 2.2 所示。由于交换机 SW1 与 SW2 不在同一个 VLAN 内，故无法通信；同理交换机 SW3 与 SW4 也不在同一个 VLAN 内，故也无法通信。虽然 SW1 与 SW3 在同一个 VLAN 内，但由于它们之间的链路处于 Discarding 状态，因此也不能通信。

为了解决这个问题，MSTP 就产生了，它可以把一台交换机的一个或多个 VLAN 划分为一个实例（Instance），有着相同 Instance 配置的交换机就组成一个区域（MST Region），运行独立的生成树（这个内部的生成树称为 IST，Internal Spanning Tree）；这个 MST Region 就相当于一台大的交换机，与其他 MST Region 再进行生成树算法运算，得出一个主干生成树，称为 CST（Common Spanning Tree）。

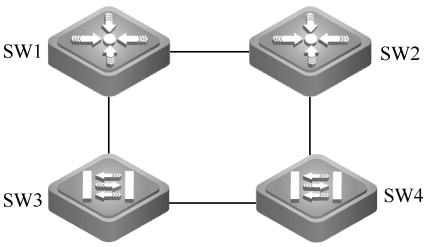


图 2.1　MSTP 多生成树协议（环路）

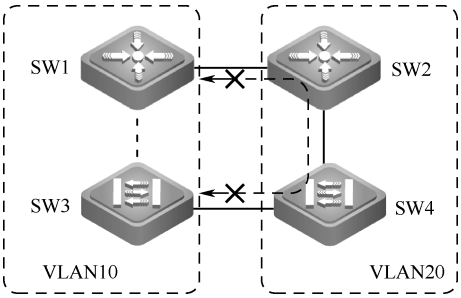


图 2.2　MSTP 多生成树协议（链路故障）

2. MSTP 的区域与实例

为了在一个或多个 VLAN 中产生一个“小树”，需要对网络中的交换机 VLAN 对实例（MST Instance）进行映射，使一个具体实例关联一个或多个 VLAN。一个 MST 实例将产生一个生成树。具有相同 MST 实例映射规则或配置的交换机组成一个 MST 区域。相同 MST 配置是指如下内容。

（1）MST 配置名称（Name）：用 32 字节长的字符串来标志 MST Region 的名称。

（2）MST Revision Number（修正号）：用 16bit 长的修正值来标志 MST Region 的修正号。

（3）VLAN 到 MST 实例的映射：在每台交换机里，最多可以创建 64 个 MST 实例（Instance），编号从 1～64，Instance 0 是强制存在的。在交换机上可以通过配置将 VLAN 和不同的 Instance 进行映射，没有被映射到 MST 实例的 VLAN 默认属于 Instance 0。实际上，在配置映射关系之前，交换机上所有的 VLAN 都属于 Instance 0。

在前面提到的网络中，SW1 和 SW3 在一个区域中，SW2 和 SW4 在另一个区域中，如图 2.3 所示。

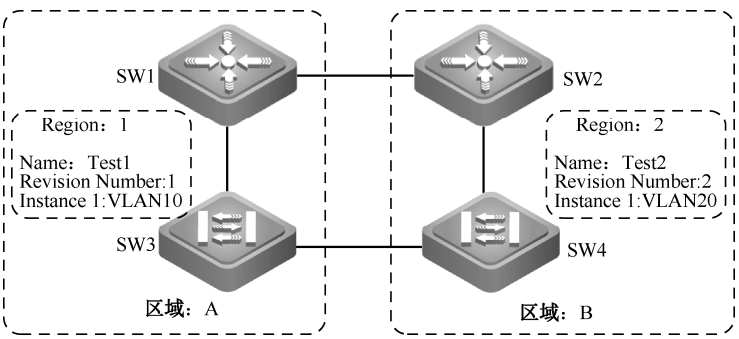


图 2.3　MSTP 多生成树协议（区域划分）

交换机 SW1 和交换机 SW3 配置相同的 MST Name：Test1，配置相同的 Revision Number：1，以及配置了相同的实例映射规则，因此在 MSTP 运算中交换机 SW1 和交换机 SW3 被认为在同一个区域。同样，交换机 SW2 和交换机 SW4 由于配置有相同的 MST Name 和 Revision Number 及相同的实例映射规则，也被视为同一区域。

3. MSTP 的多种生成树

在 MSTP 网络中，会形成很多生成树，包括 MSTI、IST、CIST、CST。每个生成树对应的范围不同，具体如下。

（1）实例生成树（MSTI）：每个实例 Instance 中的生成树叫做 MSTI（Multiple Spanning-Tree Instance）生成树，范围是该实例对应 VLAN 范围内产生的树形网络结构。

（2）区域生成树（IST）：IST（Internal Spanning Tree）是 MST 区域内的一个生成树，范围是该区域内相同配置的实例组共同产生的树形网络结构。IST 产生过程如下。

每个 Region 里按各个 Instance 所设置的 Bridge Priority、Port Priority 等参数选出各个 Instance 独立的 Root Bridge，以及每台设备上各个端口的 Port Role，然后就 Port Role 指定该端口在该 Instance 内是 Forwarding 还是 Discarding 状态。IST 使整个 MST 区域从外部上看就像一个虚拟的网桥。

（3）区域间生成树（CST）：CST（Common Spanning Tree）是连接交换网络内部的所有 MST 区域间的一个生成树。每个 MST 区域对于 CST 来说相当于一个虚拟的网桥。如果将 MST 区域视为一个网桥，那么 CST 就是这些“网桥”通过 STP 或 RSTP 计算出来的一个生成树。

（4）CIST：IST 和 CST 共同构成了整个网络的 CIST（Common and Internal Spanning Tree），它相当于每个 MST 区域中的 IST、CST 及 802.1d 网桥的集合。STP 和 RSTP 会为 CIST 选举出 CIST 的根。

按这种算法，以上网络就可以在 MSTP 算法下形成如图 2.4 所示的拓扑：交换机 SW1 和 SW3 都在 MSTP Region 1 内，MSTP Region 1 没有环路产生，所以没有链路 Discarding，同理，MSTP Region 2 的情况也是一样的。然后 Region 1 和 Region 2 就分别相当于两台大的交换机，这两台“交换机”间有环路，因此根据相关配置选择一条链路 Discarding。这样，既避免了环路的产生，也能让相同 VLAN 间的通信不受影响。

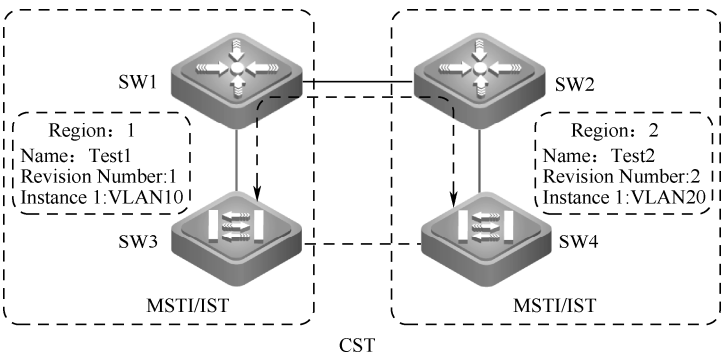


图 2.4　MSTP 多生成树协议（区域间）

MSTP 协议主要用于解决交换网络中链路负载均衡问题。通过将 VLAN 与 Instance 关联，同一台交换机对于不同 Instance 呈现出不同优先级，对于不同实例（Instance）的根（Root）交换机也不同，因此，就会形成不同的树形结构，不同流量也就会通过不同的路径，达到负载均衡作用。

如图 2.5 所示，在交换网络中 SW3 交换机划分为两个 VLAN，分别为 VLAN10 及 VLAN20。为了实现网络中数据流量负载均衡，使得 VLAN10 中的流量通过交换机 SW1 访问 Internet，VLAN20 中的流量通过交换机 SW2 访问 Internet。为了达到这一目标，启动 MSTP 多生成树协议，并且将 VLAN10 与 Instance1 关联，VLAN20 与 Instance2 关联。配置交换机 SW1 对于实例 Instance1 的优先级为 4096，对于 Instance2 为默认值 32768；配置交换机 SW2 对于实例 Instance2 的优先级为 4096，对于 Instance1 为默认值 32768，实现 SW1 为实例 Instance1 的根交换机，SW2 为实例 Instance2 的根交换机。

对于实例 Instance1，网络拓扑结构如图 2.6 所示，流量通过交换机 SW1 访问 Internet。对于实例 Instance2，网络拓扑结构如图 2.7 所示，流量通过交换机 SW2 访问 Internet。

4. MSTP 配置命令

对于 MSTP 多生成树协议配置，需要完成启动生成树协议、配置生成树类型为 MSTP、配置交换机优先级、配置端口优先级及显示 MSTP 信息等配置任务。

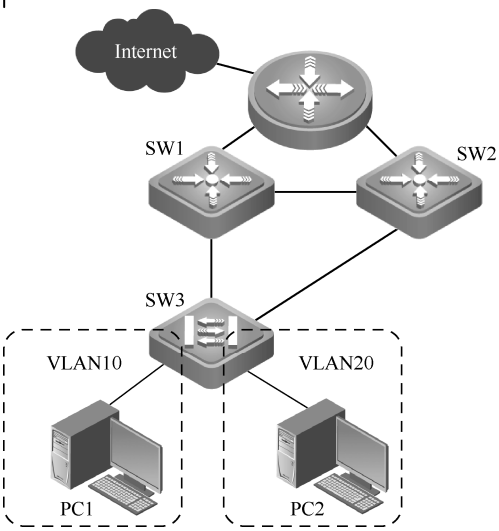


图 2.5　MSTP 多生成树协议（负载均衡）

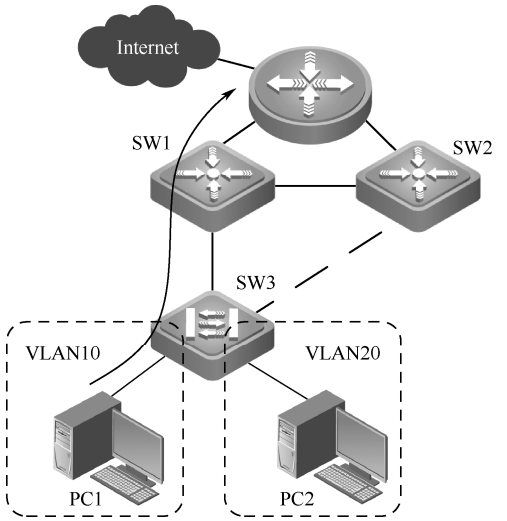


图 2.6　MSTP 多生成树协议（Instance1 流量）

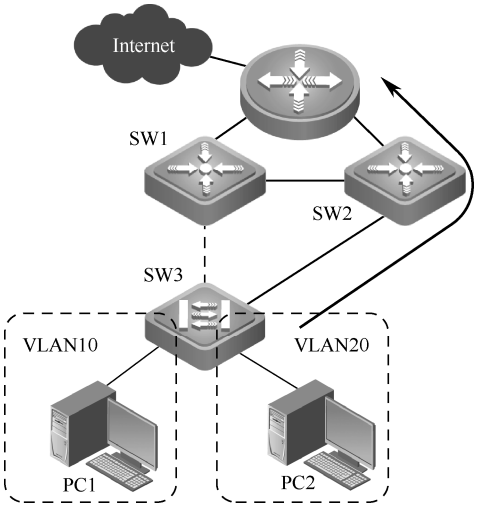


图 2.7　MSTP 多生成树协议（Instance2 流量）

（1）启动生成树协议。

启动生成树协议的命令格式如下：

启动生成树协议

ruijie (config)# spanning-tree

交换机名称为ruijie
处于全局配置模式

提示：*使用该命令的 no 选项将关闭生成树协议，如 ruijie(config)#no spanning-tree。*

例如：启动生成树协议。

```
ruijie#config
ruijie(config)#spanning-tree
```

例如：关闭生成树协议。

```
ruijie#config
ruijie(config)#no spanning-tree
```

（2）设置生成树协议类型。

交换机有 3 种类型的生成树协议：生成树协议 STP、快速生成树协议 RSTP 及多生成树协议 MSTP。默认情况下，为多生成树协议 MSTP。可以使用命令设置生成树协议类型。设置生成树协议类型的命令格式如下：

设置生成树协议类型

ruijie (config)# spanning-tree mode *mstp/rstp/stp*

交换机名称为ruijie
处于全局配置模式　　*生成树协议类型值*

其中，*mstp* 表示多生成树协议，*rstp* 表示快速生成树协议，*stp* 表示生成树协议。

提示：*要恢复默认模式，用 no spanning-tree mode 命令，如 ruijie(config)# no spanning-tree mode。*

例如：启动交换机生成树协议，并设置生成树协议类型为 MSTP。

```
ruijie#config
ruijie(config)#spanning-tree
ruijie(config)#spanning-tree mode mstp
```

例如：将交换机生成树协议类型恢复为默认状态。

```
ruijie#config
ruijie(config)#no spanning-tree mode
```

（3）配置区域属性。

为了让多台设备处于同一个 MSTP Region，需要将这几台设备的名称（Name）、版本号（Revision Number）、实例映射关系（Instance 映射 VLAN）等参数相同。具体配置步骤如下。

第 1 步：进入 MST 配置模式。

在全局配置模式下，使用 spanning-tree mst configuration 命令进入 MST 配置模式。默认情况下，Instance 和 VLAN 的对应关系是所有的 VLAN 都在 Instance 0 中；Name 为空字符串；Revision 为 0。

进入MST配置模式

ruijie(config)#spanning-tree mst configuration

交换机名称为ruijie
处于全局配置模式

提示：*使用该命令的 no 选项，将该命令下的所有参数（name、revision、VLAN map）恢复为默认值。如 ruijie(config)#no spanning-tree mst configuration。*

例如：进入 MST 配置模式。

```
ruijie(config)#spanning-tree  mst  configuration
ruijie(config-mst)#
```

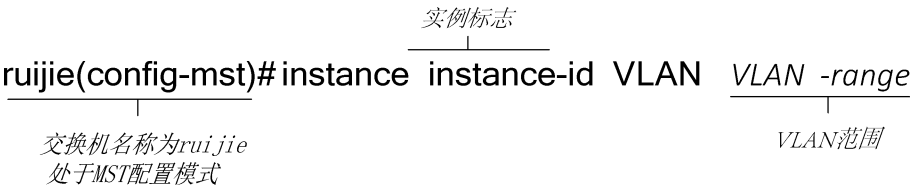
例如：恢复 MST 的默认值。

```
ruijie(config)#no spanning-tree  mst  configuration
ruijie(config-mst)#
```

第 2 步：设置 VLAN 与 Instance 关联。

进入 MST 配置模式后，需要设置 VLAN（组）与 MST Instance 的关联关系。设置关联关系的命

令格式如下：



其中，*instance-id* 为实例标志，范围为0～64；*VLAN-range* 为VLAN范围，如10～13表示VLAN10～VLAN13，15、18、19分别表示VLAN15、VLAN18、VLAN19。

提示：*使用该命令的 no 选项恢复默认值。如 ruijie(config-mst)#no instance 1。默认情况下，所有的 VLAN 均在 Instance 0 中。*

例如：将 VLAN10、VLAN20 关联到 Instance 1 上。

```
ruijie(config)#spanning-tree mst configuration
ruijie(config-mst)#instance 1 VLAN 10,20
ruijie(config-mst)#end
ruijie#show spanning-tree    mst    1
```

例如：将 VLAN20 从 Instance 1 上删除。

```
ruijie(config)#spanning-tree mst configuration
ruijie(config-mst)#no instance 1 VLAN 20
ruijie(config-mst)#end
ruijie#show spanning-tree    mst    1
```

第 3 步：配置 MST 名称。

配置 MST 名称的命令格式如下：

其中，*name* 为配置的 MST 名称，该字符串最多可以有 32 个字节，默认值为空。可以用 no name 命令将其恢复至默认值，如 ruijie(config-mst)#no name。

例如：将 MST 名称设置为 region1。

```
ruijie(config)#spanning-tree mst configuration
ruijie(config-mst)#no instance 1 VLAN 20
ruijie(config-mst)#name region1
ruijie(config-mst)#end
ruijie#show spanning-tree    mst    1
```

第 4 步：配置 MST 修正号。

配置 MST 修正号的命令格式如下：

其中，*version* 为指定的修正版本号，范围为 0～65535，默认值为 0。可以用 no revision 命令将其恢复至默认值。

例如：设置修正号为 1。

```
ruijie(config)#spanning-tree mst configuration
ruijie(config-mst)#no instance 1 VLAN 20
ruijie(config-mst)#name region1
ruijie(config-mst)#revision 1
ruijie(config-mst)#end
ruijie#show spanning-tree    mst    1
```

（4）设置实例的交换机优先级。

交换机针对不同的 MSTP 实例，可以设置为不同的优先级。在区域内针对不同的实例，分别计算该实例的根交换机、指定交换机、根端口、指定端口等参数。不同实例的根交换机不一定是同一台交换机。为了实现负载均衡，需要将不同实例的根交换机分布在不同的交换机上。设置特定实例的交换机优先级的命令格式如下：

设置MSTP交换机优先级　　*指定交换机优先级*

ruijie(config)# spanning-tree [mst *instance-id***] priority** *priority*

交换机名称为ruijie
处于全局配置模式　　*指定具体实例*

其中，*instance-id* 为实例标志，范围为 0～64；*priority* 为交换机优先级，取值范围为 0～61440，按 4096 的倍数递增，默认值为 32768。

提示：*如果要恢复到默认值，可在全局配置模式下使用 no spanning-tree mst instance-id priority 配置。*

例如：设置实例 1 的交换机优先级为 4096。

```
ruijie(config)#spanning-tree mst 1 priority 4096
ruijie(config)#
```

例如：恢复交换机的针对实例 1 的优先级默认值。

```
ruijie(config)#no spanning-tree mst 1 priority
ruijie(config)#
```

（5）配置实例的端口优先级。

配置特定实例的端口优先级的命令格式如下：

设置MSTP端口优先级　　*指定端口优先级*

ruijie(config-if)# spanning-tree mst *instance-id* **port-priority** *priority*

交换机名称为ruijie
处于端口配置模式　　*指定具体实例*

其中，*instance-id* 为实例标志，范围为 0～64；*priority* 为端口优先级，配置该 interface 的优先级，取值范围为 0～240，按 16 的倍数递增，默认值为 128。

提示：*如果要恢复到默认值，可在端口配置模式下使用 no spanning-tree mst instance-id port-priority 命令进行设置。*

例如：设置交换机 Fa0/24 端口针对实例 1 的端口优先级为 16。

```
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#spanning-tree mst 1 port-priority 16
```

例如：恢复交换机 Fa0/24 端口针对实例 1 的端口优先级默认值。

```
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#no spanning-tree mst 1 port-priority
```

（6）显示 MSTP 配置。

显示 MST 配置、实例及实例端口信息的命令格式如下：

显示MST相关信息　　*实例编号*

ruijie#show spanning-tree mst { configuration | *instance-id* [interface *interface-id*] }

交换机名称为ruijie 处于特权模式　　*MST配置信息*　　*端口编号*

其中，*configuration* 为设备的 MST 配置；*instance-id* 为 Instance 编号；*interface-id* 为端口编号。

例如：显示 MST 配置信息。

```
ruijie#show spanning-tree    mst configuration
```

默认情况下，显示所有实例的 MST 配置信息。

例如：显示实例 1 的 MST 配置信息。

```
ruijie#show spanning-tree    mst    1
```

例如：显示实例 1 的 Fa0/24 端口的 MST 配置信息。

```
ruijie#show spanning-tree    mst    1    interface    Fa 0/24
```

2.3 扩展知识

1. BPDU 概述

BPDU 是桥协议数据单元（Bridge Protocol Data Units）的简写。BPDU 的主要构成要素为 Root Bridge ID（本网桥所认为的根桥 ID）、Root Path Cost（本网桥的根路径花费）、Bridge ID（本网桥的桥 ID）、Message Age（报文已存活的时间）及 Port ID（发送该报文端口的 ID）等。

当网桥的一个端口收到高优先级的 BPDU（更小的 Bridge ID，更小的 Root Path Cost 等）时，就在该网桥保存这些信息，同时向所有端口更新并传播这些信息。如果收到比自己低优先级的 BPDU，网桥就丢弃该信息。

这样的机制就使高优先级的信息在整个网络中传播开，经过交流 BPDU，最终获得如下结果。

（1）网络中选择了一个网桥为根桥（Root Bridge）。

（2）除根桥外的每个网桥都有一个根口（Root Port），即提供最短路径到 Root Bridge 的端口。

（3）每个网桥都计算出了到根桥（Root Bridge）的最短路径。

（4）每个 LAN 都有了指派网桥（Designated Bridge），位于该 LAN 与根桥之间的最短路径中。指派网桥和 LAN 相连的端口称为指派端口（Designated Port）。

（5）根口（Root Port）和指派端口（Designated Port）进入 Forwarding 状态。

（6）其他不在生成树中的端口就处于 Discarding 状态。

2. Bridge ID 概述

按 IEEE 802.1W 标准规定，每个网桥都要有单一的网桥标志（Bridge ID），生成 MSTP 配置树算

法中就是以它为标准来选出根桥（Root Bridge）的。

Bridge ID 由 8 个字节组成，后 6 个字节为该网桥的 MAC 地址，前 2 个字节如表 2.1 所示，前 4bit 表示优先级（Priority），后 12bit 表示 System ID，为以后扩展协议而用，在 RSTP 中该值为 0，因此给网桥配置优先级就要是 4096 的倍数。

表 2.1　Bridge ID 构成

	Priority value				System ID											
Bit 位	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
值	32768	16384	8192	4096												

2.4　演示实例

1. 背景描述

某公司网络如图 2.8 所示，划分为 2 个 VLAN，分别为 VLAN10 及 VLAN20。为了使访问 Internet 的网络流量比较均衡地通过交换机 SW1 及 SW2，需要对网络中交换机进行必要配置。

为了达到网络中流量负载均衡的目标，同时实现网络链路冗余备份且不产生广播风暴，需要在网络中三台交换机中启动 MSTP 多生成树协议。创建 2 个实例 Instance1 和 Instance2，将实例 Instance1 与 VLAN10 关联，将实例 Instance2 与 VLAN20 关联。对于实例 Instance1，交换机 SW1 的优先级设置得高（如 4096），其他交换机优先级为默认值 32786；对于实例 Instance2，交换机 SW2 的优先级设置得高（如 4096），其他交换机优先级为默认值 32786。使得交换机 SW1 成为 Instance1 的根交换机，交换机 SW2 成为 Instance2 的根交换机。交换机 SW3 的端口在不同实例呈现为不同状态。

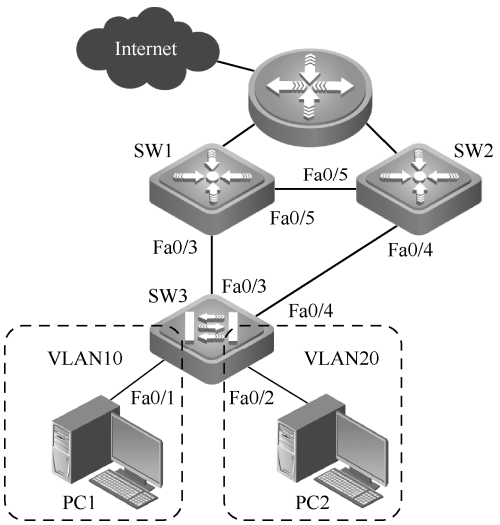


图 2.8　MSTP 操作演示

2. 操作步骤

（1）配置交换机 SW1。

第 1 步：交换机基本配置。

```
SW1(config)#VLAN 10
SW1(config-VLAN)#exit
SW1(config)#VLAN 20
SW1(config-VLAN)#exit
SW1(config)#interface FastEthernet 0/3
SW1(config-if)#Switchport mode trunk
SW1(config-if)#exit
SW1(config)#interface FastEthernet 0/5
SW1(config-if)#Switchport mode trunk
```

```
SW1(config-if)#exit
SW1(config)#
```

第 2 步：启动生成树协议。

```
SW1(config)#Spanning-tree
```

第 3 步：设置交换机生成树模式为 802.1s，即 MSTP 模式（默认状态为 MSTP）。

```
SW1(config)#Spanning-tree  mode  mstp
```

第 4 步：设置 MST 参数。

```
SW1(config)#spanning-tree mst configuration
SW1(config-mst)#revision 1
SW1(config-mst)#name test1
SW1(config-mst)#instance 1 VLAN 10
SW1(config-mst)#instance 2 VLAN 20
```

第 5 步：以核心交换机 SW1 为 Instance1 的根交换机，此交换机优先级别应为最高。将此交换机优先级设置为 4096，其他交换机为默认值 32768。

```
SW1(config)#Spanning-tree mst 1 Priority 4096
```

（2）配置交换机 SW2。

第 1 步：交换机基本配置。

```
SW2(config)#VLAN 10
SW2(config-VLAN)#exit
SW2(config)#VLAN 20
SW2(config-VLAN)#exit
SW2(config)#interface FastEthernet 0/4
SW2(config-if)#Switchport mode trunk
SW2(config-if)#exit
SW2(config)#interface FastEthernet 0/5
SW2(config-if)#Switchport mode trunk
SW2(config-if)#exit
SW2(config)#
```

第 2 步：启动生成树协议。

```
SW2(config)#Spanning-tree
```

第 3 步：设置交换机生成树模式为 802.1s，即 MSTP 模式。

```
SW2(config)#Spanning-tree  mode  mstp
```

第 4 步：设置 MST 参数。

```
SW2(config)#spanning-tree mst configuration
SW2(config-mst)#revision 1
SW2(config-mst)#name test1
SW2(config-mst)#instance 1 VLAN 10
SW2(config-mst)#instance 2 VLAN 20
```

第 5 步：以核心交换机 SW2 为 Instance2 的根交换机，此交换机优先级别应为最高，将此交换机 SW2 优先级设置为 4096，其他交换机为默认值 32768。

```
SW2(config)#Spanning-tree mst  2  Priority 4096
```

（3）配置交换机 SW3。

第 1 步：交换机基本配置。

```
SW3(config)#VLAN 10
SW3(config-VLAN)#exit
SW3(config)#VLAN 20
SW3(config-VLAN)#exit
SW3(config)#interface    FastEthernet 0/1
SW3(config-if)#Switch access VLAN 10
SW3(config-if)#exit
SW3(config)#interface    FastEthernet 0/2
SW3(config-if)#Switch access VLAN 20
SW3(config-if)#exit
SW3(config)#interface FastEthernet 0/3
SW3(config-if)#Switchport mode trunk
SW3(config-if)#exit
SW3(config)#interface FastEthernet 0/4
SW3(config-if)#Switchport mode trunk
SW3(config-if)#exit
SW3(config)#
```

第 2 步：启动生成树协议。

```
SW3(config)#Spanning-tree
```

第 3 步：设置交换机生成树模式为 802.1s，即 MSTP 模式。

```
SW3 (config)#Spanning-tree    mode    mstp
```

第 4 步：设置 MST 参数。

```
SW3(config)#spanning-tree mst configuration
SW3(config-mst)#revision 1
SW3(config-mst)#name test1
SW3(config-mst)#instance 1 VLAN 10
SW3(config-mst)#instance 2 VLAN 20
```

（4）显示 MST 配置与端口状态信息。

显示 SW1 交换机相关信息如下。

第 1 步：显示交换机 MST 配置信息。

```
SW1#show spanning-tree mst configuration
Multi spanning tree protocol : Enable
Name          : test1
Revision      : 1
Instance        VLANs Mapped
-------- ---------------------------------------------
0             : 1-9, 11-19, 21-4094
1             : 10
2             : 20
---------------------------------------------------
```

从上述显示结果可知，交换机 SW1 已启动 MSTP，MSI 名称为 test1，修正号为 1，实例 1 关联 VLAN10、实例 2 关联 VLAN20。

第 2 步：显示实例 1 的 MST 配置信息。

```
SW1#show spanning-tree mst 1
MST 1 VLANs mapped : 10
BridgeAddr : 00d0.f827.8b82
Priority: 4096
TimeSinceTopologyChange : 0d:0h:8m:23s
TopologyChanges : 2
DesignatedRoot : 1001.00d0.f827.8b82
RootCost : 0
RootPort : 0
```

从上述显示结果可知，交换机 SW1 对应实例 1 的优先级为 4096，SW1 为实例 1 的根交换机。

第 3 步：显示针对实例 1 的 Fa0/3 端口状态信息。

```
SW1#show spanning-tree mst 1 interface fa 0/3
MST 1 VLANs mapped :10
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1001.00d0.f827.8b82
PortDesignatedCost : 0
PortDesignatedBridge : 1001.00d0.f827.8b82
PortDesignatedPort : 8003
PortForwardTransitions : 2
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/3 端口对于实例 1 来讲，为指定端口，处于 forwarding 状态。

第 4 步：显示针对实例 1 的 Fa0/5 端口状态。

```
SW1#sh spanning-tree mst 1 interface fa 0/5
MST 1 VLANs mapped :10
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1001.00d0.f827.8b82
PortDesignatedCost : 0
PortDesignatedBridge : 1001.00d0.f827.8b82
PortDesignatedPort : 8005
PortForwardTransitions : 1
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/5 端口对于实例 1 来讲，为指定端口，处于 forwarding 状态。

第 5 步：显示实例 2 的 MST 配置信息。

```
SW1#show spanning-tree mst 2
MST 2 VLANs mapped : 20
BridgeAddr : 00d0.f827.8b82
Priority: 32768
```

```
TimeSinceTopologyChange : 0d:0h:8m:28s
TopologyChanges : 3
DesignatedRoot : 1002.00d0.f827.b202
RootCost : 200000
RootPort : 5
```

从上述显示结果可知，交换机 SW1 对应实例 2 的优先级为 32768，SW1 为实例 2 的指定交换机，SW1 上的 Fa0/5 为实例 2 的根端口。

第 6 步：显示针对实例 2 的 Fa0/3 端口状态信息。

```
SW1#show spanning-tree mst 2 interface fa 0/3
MST 2 VLANs mapped :20
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1002.00d0.f827.b202
PortDesignatedCost : 0
PortDesignatedBridge : 8002.00d0.f827.8b82
PortDesignatedPort : 8003
PortForwardTransitions : 2
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/3 端口对于实例 2 来讲，为指定端口，处于 forwarding 状态。

第 7 步：显示针对实例 2 的 Fa0/5 端口状态信息。

```
SW1#show spanning-tree mst 2 interface fa 0/5
MST 2 VLANs mapped :20
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1002.00d0.f827.b202
PortDesignatedCost : 0
PortDesignatedBridge : 1002.00d0.f827.b202
PortDesignatedPort : 8005
PortForwardTransitions : 1
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : rootPort
```

从上述显示结果可知，Fa0/5 端口对于实例 2 来讲，为根端口，处于 forwarding 状态。

显示 SW2 交换机相关信息如下。

第 1 步：显示交换机 MST 配置信息。

```
SW2#show spanning-tree mst configuration
Multi spanning tree protocol : Enable
Name        : test1
Revision    : 1
Instance    VLANs Mapped
-------- ---------------------------------------------
```

```
0          : 1-9, 11-19, 21-4094
1          : 10
2          : 20
-------------------------------------------------------
```

从上述显示结果可知，交换机 SW2 已启动 MSTP，MSI 名称为 test1，修正号为 1，实例 1 关联 VLAN10、实例 2 关联 VLAN20。

第 2 步：显示实例 1 的 MST 配置信息。

```
SW2#show spanning-tree mst 1
MST 1 VLANs mapped : 10
BridgeAddr : 00d0.f827.b202
Priority: 32768
TimeSinceTopologyChange : 0d:0h:9m:55s
TopologyChanges : 3
DesignatedRoot : 1001.00d0.f827.8b82
RootCost : 200000
RootPort : 5
```

从上述显示结果可知，交换机 SW2 对应实例 1 的优先级为 32768，SW2 为实例 1 的指定交换机。

第 3 步：显示针对实例 1 的 Fa0/4 端口状态信息。

```
SW2#show spanning-tree mst 1 interface fastEthernet 0/4
MST 1 VLANs mapped :10
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1001.00d0.f827.8b82
PortDesignatedCost : 0
PortDesignatedBridge : 8001.00d0.f827.b202
PortDesignatedPort : 8004
PortForwardTransitions : 3
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/4 端口对于实例 1 来讲，为指定端口，处于 forwarding 状态。

第 4 步：显示针对实例 1 的 Fa0/5 端口状态信息。

```
SW2#show spanning-tree mst 1 interface fastEthernet 0/5
MST 1 VLANs mapped :10
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1001.00d0.f827.8b82
PortDesignatedCost : 0
PortDesignatedBridge : 1001.00d0.f827.8b82
PortDesignatedPort : 8005
PortForwardTransitions : 1
PortAdminPathCost : 0
PortoperPathCost : 200000
```

PortRole : **rootPort**

从上述显示结果可知，Fa0/5 端口对于实例 1 来讲，为根端口，处于 forwarding 状态。

第 5 步：显示实例 2 的 MST 配置信息。

```
SW2#show spanning-tree mst 2
MST 2 VLANs mapped : 20
BridgeAddr : 00d0.f827.b202
Priority: 4096
TimeSinceTopologyChange : 0d:0h:9m:58s
TopologyChanges : 2
DesignatedRoot : 1002.00d0.f827.b202
RootCost : 0
RootPort : 0
```

从上述显示结果可知，交换机 SW2 对应实例 2 的优先级为 4096，SW2 为实例 2 的根交换机。

第 6 步：显示针对实例 2 的 Fa0/4 端口状态信息。

```
SW2#show spanning-tree mst 2 interface fastEthernet 0/4
MST 2 VLANs mapped :20
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1002.00d0.f827.b202
PortDesignatedCost : 0
PortDesignatedBridge : 1002.00d0.f827.b202
PortDesignatedPort : 8004
PortForwardTransitions : 2
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/4 端口对于实例 2 来讲，为指定端口，处于 forwarding 状态。

第 7 步：显示针对实例 2 的 Fa0/5 端口状态信息。

```
SW2#show spanning-tree mst 2 interface fastEthernet 0/5
MST 2 VLANs mapped :20
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 1002.00d0.f827.b202
PortDesignatedCost : 0
PortDesignatedBridge : 1002.00d0.f827.b202
PortDesignatedPort : 8005
PortForwardTransitions : 1
PortAdminPathCost : 0
PortoperPathCost : 200000
PortRole : designatedPort
```

从上述显示结果可知，Fa0/5 端口对于实例 2 来讲，为指定端口，处于 forwarding 状态。

显示 SW3 交换机相关信息如下。

第 1 步：显示交换机 MST 配置信息。

```
SW3#show spanning-tree mst configuration
Multi spanning tree protocol : Enabled
Name          : test1
Revision      : 1
Instance      VLANs Mapped
--------    ------------------------------------------------------------
0             1-9,11-19,21-4094
1             10
2             20
---------------------------------------------------------------------
```

从上述显示结果可知，交换机 SW3 已启动 MSTP，MSI 名称为 test1，修正号为 1，实例 1 关联 VLAN10、实例 2 关联 VLAN20。

第 2 步：显示实例 1 的 MST 配置信息。

```
SW3#show spanning-tree mst 1
MST 1 VLANs mapped : 10
BridgeAddr : 00d0.f8ef.b043
Priority : 32768
TimeSinceTopologyChange : 0d:1h:41m:35s
TopologyChanges : 0
DesignatedRoot : 100100D0F8278B82
RootCost : 200000
RootPort : Fa0/3
```

从上述显示结果可知，交换机 SW3 对应实例 1 的优先级为 32768，SW3 为实例 1 的指定交换机，Fa0/3 为根端口。

第 3 步：显示针对实例 1 的 Fa0/3 端口状态。

```
SW3#show spanning-tree mst 1 interface fastEthernet 0/3
MST 1 VLANs mapped : 10
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 100100D0F8278B82
PortDesignatedCost : 0
PortDesignatedBridge : 100100D0F8278B82
PortDesignatedPort : 8003
PortForwardTransitions : 2
PortAdminPathCost : 0
PortOperPathCost : 200000
PortRole : rootPort
```

从上述显示结果可知，Fa0/3 端口对于实例 1，为根端口，处于 forwarding 状态。

第 4 步：显示针对实例 1 的 Fa0/4 端口状态。

```
SW3#show spanning-tree mst 1 interface fastEthernet 0/4
MST 1 VLANs mapped : 10
PortState : discarding
PortPriority : 128
```

PortDesignatedRoot : 100100D0F8278B82
PortDesignatedCost : 200000
PortDesignatedBridge : 800100D0F827B202
PortDesignatedPort : 8004
PortForwardTransitions : 1
PortAdminPathCost : 0
PortOperPathCost : 200000
PortRole : **alternatePort**

从上述显示结果可知，Fa0/4 端口对于实例 1，为替换端口，作为根端口的备份端口，处于 discarding 状态。

第 5 步：显示实例 2 的 MST 配置信息。

SW3#show spanning-tree mst 2
MST 2 VLANs mapped : 20
BridgeAddr : 00d0.f8ef.b043
Priority : **32768**
TimeSinceTopologyChange : 0d:1h:41m:38s
TopologyChanges : 0
DesignatedRoot : 100200D0F827B202
RootCost : 200000
RootPort : **Fa0/4**

从上述显示结果可知，交换机 SW3 对应实例 2 的优先级为 32768，SW3 为实例 2 的指定交换机，Fa0/4 为根端口。

第 6 步：显示针对实例 2 的 Fa0/3 端口状态。

SW3#show spanning-tree mst 2 interface fastEthernet 0/3
MST 2 VLANs mapped : 20
PortState : **discarding**
PortPriority : 128
PortDesignatedRoot : 100200D0F827B202
PortDesignatedCost : 200000
PortDesignatedBridge : 800200D0F8278B82
PortDesignatedPort : 8003
PortForwardTransitions : 2
PortAdminPathCost : 0
PortOperPathCost : 200000
PortRole : **alternatePort**

从上述显示结果可知，Fa0/3 端口对于实例 2，为替换端口，作为根端口的备份端口，处于 discarding 状态。

第 7 步：显示针对实例 2 的 Fa0/4 端口状态。

SW3#show spanning-tree mst 2 interface fastEthernet 0/4
MST 2 VLANs mapped : 20
PortState : **forwarding**
PortPriority : 128
PortDesignatedRoot : 100200D0F827B202

```
PortDesignatedCost : 0
PortDesignatedBridge : 100200D0F827B202
PortDesignatedPort : 8004
PortForwardTransitions : 1
PortAdminPathCost : 0
PortOperPathCost : 200000
PortRole : rootPort
```

从上述显示结果可知，Fa0/4 端口对于实例 2，为根端口，处于 forwarding 状态。

提示：*网络链路状态发生变化后，交换机生成树协议需要时间进行重新调整，需要稍等几秒钟时间才能观察到切换后的结果。*

2.5 训练任务

1. 背景描述

某小型办公区域网络中有计算机 5 台，汇聚层交换机 S3760 2 台、接入层交换机 S2126G 2 台，需要将办公区域内的计算机划分成 4 个区域，分别是 VLAN10、VLAN20、VLAN30 及 VLAN40，接入层交换机每个 VLAN 连接一台计算机，通过配置交换机参数，到达 SW1 成为 VLAN10、VLAN20 的根交换机，SW2 成为 VLAN30、VLAN40 的根交换机，实现网络链路负载均衡。参考拓扑图如图 2.9 所示。

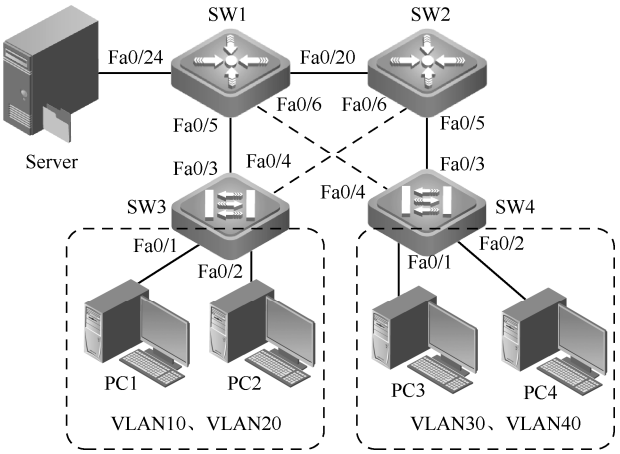


图 2.9　MSTP 训练网络拓扑图

2. 操作提示

在配置 MSTP 时需要主要以下几点。

（1）SW1、SW2、SW3 及 SW4 都要分别启动 MSTP 协议，并进行相应配置。

（2）对于 SW1，VLAN10、VLAN20 的实例优先级为最高，SW1 为其实例的根交换机。同样，对于 SW2，VLAN30、VLAN40 的实例优先级为最高，SW2 为其实例的根交换机。

练习题

1. 选择题

（1）定义 MSTP 的是（　　）。

A．802.1d　　B．802.1q　　C．802.1w　　D．802.1s

（2）MSTP 要处理以下（　　）MST 实例。

A．CIST　　B．IST　　C．MST　　D．MSTI

（3）具有（　　）MST 实例映射规则的交换机组成一个 MST 区域。

A．MST 配置名称（Name）　　B．MST Revision Number（修正号）

C．优先级　　D．VLAN 到 MST 实例的映射

2. 简答题

（1）描述 MSTP 做了哪些改进，实现哪些功能。

（2）比较 STP、RSTP、MSTP 的区别，说明它们的异同点。

（3）描述 MSTP 拓扑发生改变时的工作过程。

项目 3　虚拟路由器冗余协议 VRRP

当数据包跨网段转发时，必须通过网关。如果网关出现宕机状况，则数据包就无法通过网关到达其他网络。如何确保网络中有可靠的网关在工作，对于保障网络安全、可靠地传输数据十分必要。使用 VRRP 协议可以解决此问题，同时，还可以提供负载均衡解决方案。

通过本章学习能够对交换机 VRRP 协议有所了解，掌握交换机 VRRP 协议特性及配置技能。

知识点、技能点

1. 了解 VRRP 协议;
2. 掌握 VRRP 特性及配置技能;
3. 掌握小型网络一般调试技能及故障排除方法。

3.1　问题提出

在 TCP/IP 网络结构中，为了实现 IP 数据包跨网段访问，网络设备一般需设置默认网关 IP 地址，在查询路由表未果的情况下通过默认网关设备转发数据包。如图 3.1 所示网络，在 Ethernet 中所有 PC 的默认网关 IP 地址设置为路由器 Router1 的 Fa0/0 接口的 IP 地址 192.168.1.1。PC 如果要访问外部网络，将数据包转发给路由器 Router1，通过路由器实现访问其他网络的目的。

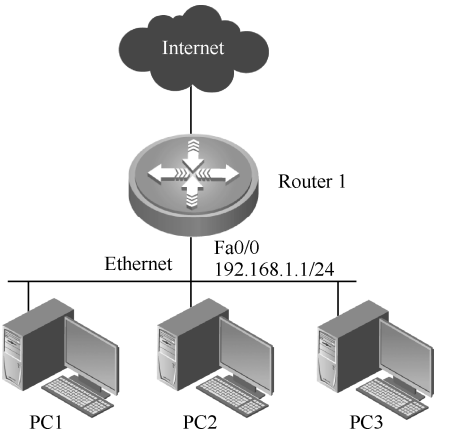


图 3.1　配置默认网关

但是，如果作为默认网关的路由器出现故障，所有使用该路由器作为默认网关的主机必然因网关故障而通信中断。即便配置了多个默认网关，如不重新启动终端设备，也不能切换到新的网关。那么，采取什么措施才能有效解决这样的问题呢？

一种解决方案是在网络上多部署一台路由器，为主机配置多个默认网关。这种解决方案初看起来实现了网关冗余，如果第 1 个默认网关失效，就通过第 2 个默认网关转发数据，但是，从实践运行情况看并不能真正地做到网关冗余。如图 3.2 所示，PC1 设置双网关（192.168.1.1 及 192.168.1.2）去访

问 Internet，如果 Router1 正常工作，那么 PC1 就会把数据包发送给网关地址 192.168.1.1，然后由 Router1 路由给 Internet。如果 Router 1 因某种原因链路 down 了，那么 PC1 是无法感知到这个故障的，而继续将报文发送给网关地址 192.168.1.1，也就是说，在这种情况下，没有什么机制可以使 PC1 切换到另一个网关地址 192.168.1.2。

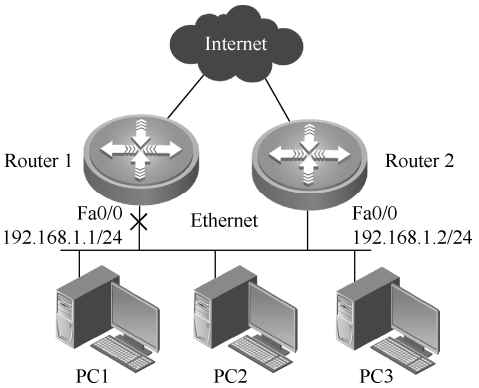


图 3.2　配置多个默认网关

另一种解决方案是在两台路由器上启动 VRRP，当 Router1 工作正常时，由其转发数据；当 Router1 故障时，系统自动切换至 Router2，由其代替 Router1 实现无间断转发。

3.2　相关知识

1．VRRP 概述

VRRP（Virtual Router Redundancy Protocol，虚拟路由器冗余协议）是一种网关备份冗余解决方案，它对 Ethernet 上的计算机等终端的默认网关进行冗余备份，实现当其中一台网关设备宕机时，备份网关设备能够及时自动切换并接管转发工作，防止因网关设备故障引发转发数据失败，提高网络服务质量。

VRRP 通过使用虚拟路由器技术实现了主机默认网关的备份，同时也可以通过 VRRP 来实现网关的负载均衡。

在 VRRP 中通过设置将路由器分成若干组，并将每个分组分配一个编号，用于标志这个组。每一个组中可以包含两台或多台路由器，每台路由器也可以属于一个或多个组。每一个组分配一个 IP 地址及 MAC 地址。常常将这个路由器组称为虚拟路由器（Virtual Router），如图 3.3 所示。

在路由器组中的路由器处于两种角色，分别是主路由器和备份路由器。主路由器负责 ARP 响应和转发 IP 数据包；组中的其他路由器作为备份的角色处于待命状态。VRRP 使用选举机制从一组 VRRP 路由器中选出一台作为主路由器，一个 VRRP 组中有且只有一台处于主控角色的路由器，可以有一台或者多台处于备份角色的路由器。当由于某种原因主路由器发生故障时，备份路由器能在几秒钟的时延后升级为主路由器。由于切换非常迅速而且终端不用改变默认网关的 IP 地址和 MAC 地址，故对终端使用者来说系统是透明的。

在图 3.3 中，由 Router1、Router2 构造一台 VRRP 虚拟路由器，这台虚拟路由器的 IP 地址被设置为 Router1 的 IP 地址 192.168.1.254，网络中所有主机的默认网关都为虚拟路由器的 IP 地址。

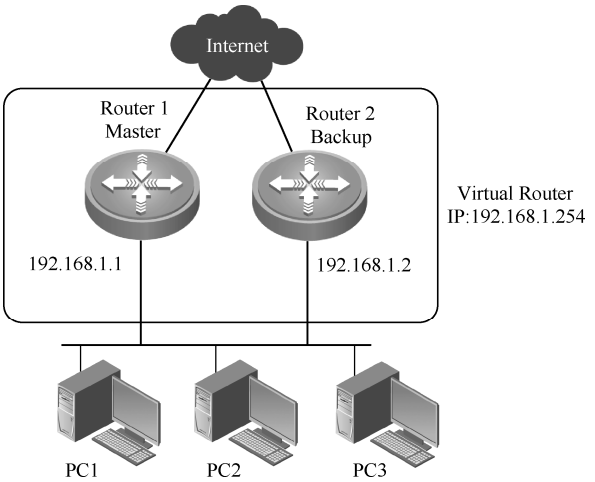


图 3.3　虚拟路由器组

由于虚拟路由器使用 Router1 的物理以太网接口的 IP 地址，因此 Router1 就担当了主路由器的角色，Router1 被称为 IP 地址拥有者（owner）。作为主路由器，Router1 控制虚拟路由器的 IP 地址，并负责对发送到该虚拟 IP 地址的数据包进行转发。Router2 为备用路由器，如果主路由器 Router1 发生故障，Router2 作为备用路由器将替代为主路由器。当 Router1 恢复正常后，将再次成为主路由器。

一个 VRRP 组中的路由器都有唯一的标志——VRID，范围为 0～255，它决定运行 VRRP 的路由器是属于哪一个 VRRP 组的。VRRP 组中的虚拟路由器对外表现为唯一的虚拟 MAC 地址，地址的格式为 00-00-5E-00-01-[VRID]。主路由器负责对发送到虚拟路由器 IP 地址的 ARP 请求做响应并以该虚拟 MAC 地址做应答。这样，无论如何切换，保证给终端设备的是唯一一致的 IP 地址和 MAC 地址，避免了切换对终端设备的影响。

2．VRRP 状态

VRRP 路由器在运行过程中会有三种状态，分别是 Initialize、Master 和 Backup。

（1）Initialize 状态。系统启动后进入 Initialize 状态，在此状态时，路由器不会对 VRRP 报文做任何处理。当收到接口 Up 的消息后，将进入 Backup 状态或 Master 状态。

（2）Master 状态。当路由器处于 Master 状态时，它将执行以下任务。

- 发送免费 ARP（Gratuitous ARP）报文，以使网络内各主机知道虚拟 IP 地址所对应的虚拟 MAC 地址。
- 响应对虚拟 IP 地址的 ARP 请求，并且响应的是虚拟 MAC 地址，而不是接口的真实 MAC 地址。
- 转发目的 MAC 地址为虚拟 MAC 地址的 IP 报文。
- 如果它是这个虚拟 IP 地址的拥有者，则接收目的 IP 地址为这个虚拟 IP 地址的 IP 报文，否则丢弃这个 IP 报文。

在 Master 状态中，只有当接收到接口的 shutdown 事件时才会转为 Initialize 状态。

（3）Backup 状态。当路由器处于 Backup 状态时，它将执行以下任务。

- 接收 Master 发送的 VRRP 报文，从中了解 Master 的状态。
- 对虚拟 IP 地址的 ARP 请求不做响应。
- 丢弃目的 MAC 地址为虚拟 MAC 地址的 IP 报文。

■ 丢弃目的 IP 地址为虚拟 IP 地址的 IP 报文。

同样，在 Backup 状态中，只有当接收到接口的 shutdown 事件时才会转为 Initialize 状态。

3．VRRP 选举机制

VRRP 使用选举机制来确定路由器的状态（Master 或 Backup）。VRRP 路由器通过交换 VRRP 通告消息，获知其他路由器的 VRRP 优先级信息。通过比较路由器的优先级进行选举，优先级高的路由器将成为主路由器，其他路由器都为备份路由器。

如果 VRRP 组中存在 IP 地址拥有者，即虚拟 IP 地址与某台 VRRP 路由器的地址相同时，IP 地址拥有者将成为主路由器，并且 IP 地址拥有者将具有最高的优先级 255。如果 VRRP 组中不存在 IP 地址拥有者，VRRP 路由器将通过比较优先级来确定主路由器。默认情况下，VRRP 路由器的优先级为 100。当优先级相同时，VRRP 将通过比较 IP 地址，IP 地址大的路由器将成为主路由器。

如图 3.4 所示，Router1 的 VRRP 优先级为 120，Router2 的 VRRP 优先级为默认的 100，那么 Router1 将成为该组的主路由器（Master），Router2 成为该组的备份路由器（Backup）。

如果路由器 Router1 和 Router2 优先级相同，将通过比较接口的 IP 地址，接口 IP 地址大者优先成为主路由器。假设在图 3.4 中，如果两台路由器都没有设置优先级，即优先级都是默认值 100，比较接口 IP 地址，Router2 成为主路由器，Router1 成为备份路由器。

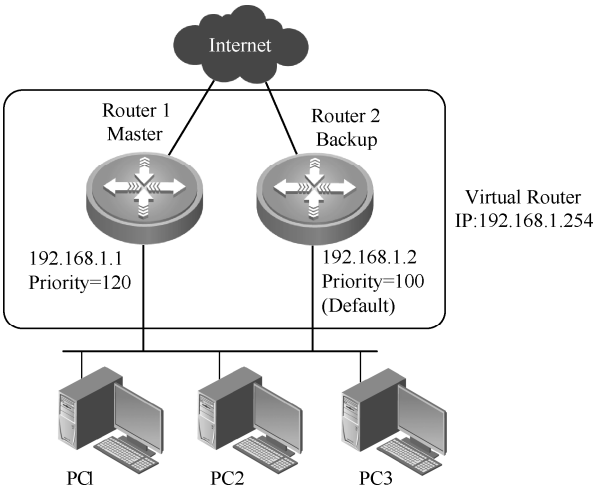


图 3.4　VRRP 选举

4．VRRP 定时器

VRRP 在运行过程中使用两个定时器用来进行状态检测。

（1）通告定时器（adver-timer）：该定时器在主路由器中使用，用来定义通告间隔（adver-interval）。主路由器以该定时器的间隔定期发送 VRRP 通告报文，用来告知其他备份路由器自己仍在线。通告间隔默认为 1 秒，可以通过配置进行修改。

（2）主路由器失效定时器（master-down-timer）：该定时器在备份路由器中使用，用来定义主路由器失效间隔（master-down-interval）。主路由器失效间隔指的是备用路由器多长时间没有收到主路由器的通告报文后，将认为主路由器已失效，并开始选举新的主路由器。主路由器失效间隔是通告间隔的 3 倍，默认为 3 秒。

5. VRRP 基本配置

（1）创建 VRRP 组，并配置虚拟 IP 地址。要启用 VRRP 并使 VRRP 能够正常工作，最基本的配置是要创建 VRRP 组，并为 VRRP 组配置虚拟 IP 地址。

在接口模式下，使用如下命令创建 VRRP 组，并配置虚拟 IP 地址：

创建的VRRP组　　*设置虚拟组次IP地址*

ruijie(config-if)# vrrp *group* **ip** *ip-address* **[secondary]**

交换机名称为ruijie
处于接口配置模式　　*虚拟组IP地址*

其中，*group* 为 VRRP 组的编号，即 VRID，取值范围为 1～255。属于同一个 VRRP 组的路由器必须配置相同的 VRID 才能正常工作。对于一台路由器将其加入到多个 VRRP 组中。

ip-address 为 VRRP 组的虚拟 IP 地址。虚拟 IP 地址可以是该子网中未使用的地址，也可以是某台 VRRP 路由器的接口 IP 地址，即 IP 地址拥有者。虚拟 IP 地址必须与接口地址位于同一个子网中。

secondary 为该 VRRP 组配置辅助 IP 地址。

默认情况下，接口没有启用 VRRP 功能，即没有 VRRP 组。使用该命令的 no 选项，将禁止该接口上的 VRRP 功能并取消虚拟 IP 地址的设置。如 ruijie(config-if)#no vrrp 10 ip 192.168.10.1。

例如：在如图 3.5 所示的网络拓扑中，Router1 与 Router2 属于 VRRP 组 10，虚拟 IP 地址为 Router1 接口的地址，所以 Router1 成为该组的 IP 地址拥有者和主路由器。所有主机的默认网关设置为虚拟 IP 地址。

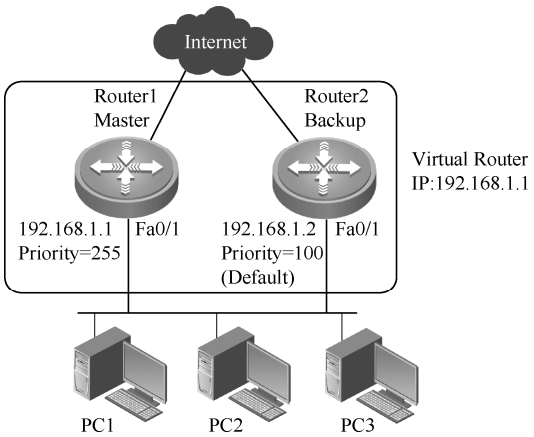


图 3.5　配置 VRRP 组

配置代码如下。

第 1 步：Router1 的 VRRP 组配置。

```
Router1(config)#interface FastEthernet 0/1
Router1(config-if)#ip address 192.168.1.1    255.255.255.0
Router1(config-if)#vrrp 10    ip 192.168.1.1
Router1(config-if)#end
```

第 2 步：Router2 的 VRRP 组配置。

```
Router2(config)#interface FastEthernet 0/1
Router2(config-if)#ip address 192.168.1.2    255.255.255.0
```

```
Router2(config-if)#vrrp 10    ip 192.168.1.1
Router2(config-if)#end
```

配置完成后，可以使用命令 show vrrp brief 来查看 VRRP 组的状态。

第 3 步：查看 Router1 的 VRRP 状态。

```
Router1#show vrrp brief
Interface          Grp  Pri  Time  Own  Pre  State    Master addr     Group addr
FastEthernet 1/0   10   255  3     O    P    Master   192.168.1.1     192.168.1.1
```

从 Router1 的显示信息中可以看出，Router1 的优先级为 255，作为 VRRP 组 10 的 IP 地址拥有者，并且状态为 Master（主路由器）。

第 4 步：查看 Router2 的 VRRP 状态。

```
Router2#show vrrp brief
Interface          Grp  Pri  Time  Own  Pre  State    Master addr     Group addr
FastEthernet 1/0   10   100  3     -    P    Backup   192.168.1.2     192.168.1.1
```

从 Router2 的显示信息中可以看出，Router2 的优先级为默认值 100，状态为 Backup（备份路由器）。

（2）配置 VRRP 优先级。VRRP 通过比较优先级来选举主路由器和备份路由器。如果 VRRP 组中存在 IP 地址拥有者，那么其优先级为最高数值 255，并成为主路由器。如果 VRRP 组中不存在 IP 地址拥有者，即虚拟 IP 地址不与任何路由器接口地址相同，就需要通过比较优先级来选举。

默认情况下，VRRP 路由器的优先级为 100。在优先级相同的情况下，IP 地址大的将成为主路由器。如果我们希望某台路由器成为主路由器，可以手工调整其优先级来影响选举结果。

如图 3.6 所示，我们使用两台路由器 Router1 及 Router2 作为 LAN1 双出口连接到外部网络，并希望高带宽的链路作为主要链路，低带宽链路作为主链路的备份。这种情况下，我们可以手工调整路由器的 VRRP 优先级，为连接主链路的路由器 Router1 配置更高的优先级，使其成为主路由器。此外，我们还可以根据路由器的性能来调整优先级，为性能好的路由器配置更高的优先级。

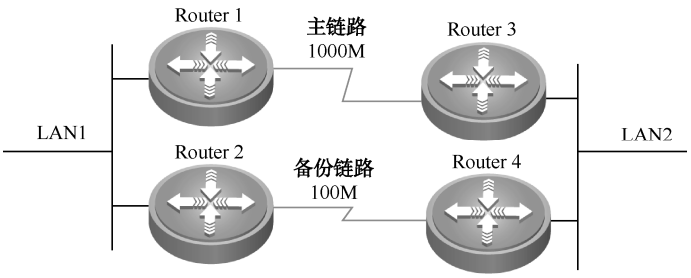


图 3.6　配置 VRRP 优先级

优先级的配置是基于接口和 VRRP 组的，也就是说，对于不同的接口和不同的 VRRP 组可以分配不同的优先级值。如图 3.7 所示网络，在虚拟组 10 中 Router1 路由器的 Fa0/0 接口的优先级为 120，Router2 路由器的 Fa0/0 接口的优先级为默认值 100，Router1 成为主路由器，Router2 成为备份路由器；在虚拟组 20 中 Router1 路由器的 Fa0/0 接口的优先级为默认值 100，Router2 路由器的 Fa0/0 接口的优先级为 120，Router1 成为备份路由器，Router2 成为主路由器。

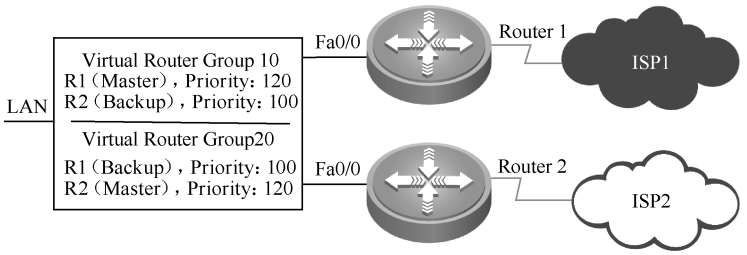


图 3.7　VRRP 基于接口配置优先级

在接口配置模式下，使用如下命令修改默认的优先级：

创建的VRRP组

ruijie(config-if)# vrrp *group* **priority** *level*

交换机名称为ruijie
处于接口配置模式　　*设置优先级*

其中，*group* 表示 VRRP 组号；*level* 表示优先级，取值范围为 1～254，默认为 100。实际上，VRRP 的优先级的范围为 0～255，0 被保留为特殊用途使用，255 表示 IP 地址拥有者。

使用该命令的 no 选项，将恢复系统的默认值，如 ruijie(config-if)#no vrrp 10 priority。

例如：对如图 3.7 所示网络，进行设备 VRRP 配置，实现负载均衡。配置代码如下。

第 1 步：在 Router1 上创建 VRRP 组 10，设置 IP 地址为 192.168.1.253，VRRP 优先级为 120；创建 VRRP 组 20，设置 IP 地址为 192.168.1.254，VRRP 优先级为默认值 100。

```
Router1(config)#interface FastEthernet 0/0
Router1(config-if)#ip address 192.168.1.1    255.255.255.0
Router1(config-if)#vrrp 10    ip 192.168.1.253
Router1(config-if)#vrrp 10    priority 120
Router1(config-if)#vrrp 20    ip 192.168.1.254
Router1(config-if)#end
```

第 2 步：在 Router2 上，创建 VRRP 组 10，设置 IP 地址为 192.168.1.253，VRRP 优先级为默认值 100；创建 VRRP 组 20，设置 IP 地址为 192.168.1.254，VRRP 优先级为 120。

```
Router2(config)#interface FastEthernet 0/0
Router2(config-if)#ip address 192.168.1.2    255.255.255.0
Router2(config-if)#vrrp 10    ip 192.168.1.253
Router2(config-if)#vrrp 20    ip 192.168.1.254
Router2(config-if)#vrrp 20    priority 120
Router2(config-if)#end
```

配置完成后，可以使用命令 show vrrp brief 来查看 VRRP 组的状态。

第 3 步：查看 Router1 的 VRRP 状态。

```
Router1#show vrrp brief
Interface              Grp   Pri   Time   Own Pre   State     Master addr      Group addr
FastEthernet 0/0       10    120    3       -       P       Master   192.168.1.1     192.168.1.253
FastEthernet 0/0       20    100    3       -       P       Bakup    192.168.1.2     192.168.1.254
```

对于 Router1 的 Fa0/0，在 Group10 中优先级为 120，路由器为 Master（主路由器）；在 Group20 中优先级为默认值 100，路由器为 Backup（备份路由器）。

第 4 步：查看 Router 2 的 VRRP 状态。

```
router2#show vrrp brief
Interface          Grp  Pri  Time  Own  Pre  State    Master addr     Group addr
FastEthernet 0/0   10   100  3     -    P    Backup   192.168.1.1     192.168.1.253
FastEthernet 0/0   20   120  3     -    P    Master   192.168.1.2     192.168.1.25
```

对于 Router2 的 Fa0/0，在 Group10 中优先级为默认值 100，路由器为 Backup（备份路由器）；在 Group20 中优先级为默认值 120，路由器为 Master（主路由器）。

6．VRRP 负载均衡配置

在标准的 VRRP 运行环境中，主路由器负责转发到达虚拟 IP 地址的数据，备份路由器不负责数据的转发，它只是侦听主路由器的状态，以在必要的时刻进行故障切换。但在主路由器承担数据转发任务的同时，备份路由器的链路将处于空闲状态，势必造成了带宽资源的浪费。在如图 3.8 所示的网络拓扑中，Router1 的 100M 链路的带宽将被浪费。

为了能够提高冗余性，并且避免造成带宽资源的浪费，我们可以在 VRRP 中使用负载均衡。VRRP 负载均衡是通过将路由器加入到多个 VRRP 组实现的，使 VRRP 路由器在不同的组中担任不同的角色，如图 3.8 所示。

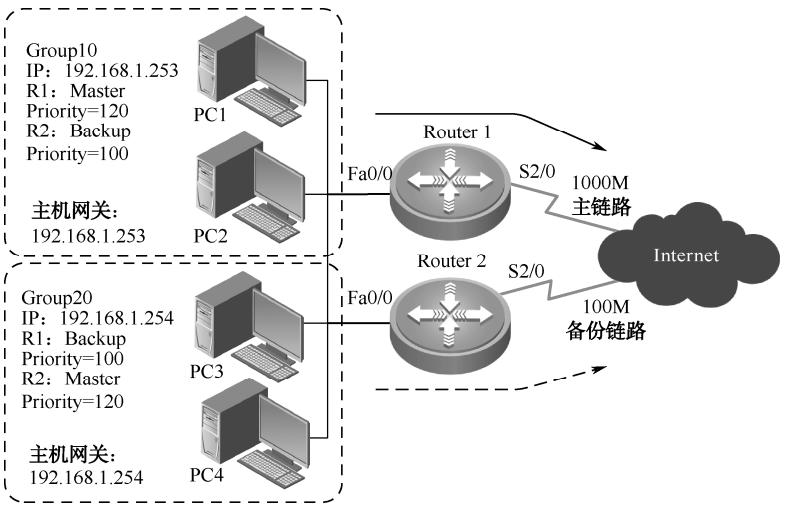


图 3.8　VRRP 负载均衡

在图 3.8 中，Router1 和 Router2 的 Fa0/0 接口都加入到了 VRRP 组 10 和 VRRP 组 20。在 VRRP 组 10 中 Router1 的优先级为 120，Router2 的优先级为默认值 100，Router1 成为主路由器，Router2 成为备份路由器；在 VRRP 组 20 中 Router1 的优先级为默认值 100，Router2 的优先级为 120，Router2 成为主路由器，Router1 成为备份路由器。VRRP 组 10 的虚拟地址为 192.168.1.253，VRRP 组 20 的虚拟地址为 192.168.1.254。在主机的配置中，PC1 和 PC2 的默认网关为 VRRP 组 10 的虚拟地址 192.168.1.253，PC3 和 PC4 的默认网关为 VRRP 组 20 的虚拟地址 192.168.1.254。

通过这样的部署和配置，我们可以看到，PC1 和 PC2 发送到其他子网的数据流由 Router1 转发，PC3 和 PC4 发送到其他子网的数据流由 Router2 转发。这样 Router1 和 Router2 带宽都被合理地利用，避免了某条链路由于作为备份而产生的空闲状态。当在图 3.8 中的拓扑使用 VRRP 负载均衡时，Router1 和 Route2 的两条链路都能够被有效地利用，不仅提高了冗余性，还提供了流量的负载均衡。

实际上，VRRP 并不具备对流量进行监控的机制，它的负载均衡是通过使用多个 VRRP 组来实现的，并且这种负载均衡还需要终端配置的配合，即让不同的终端将数据发送到不同的 VRRP 组 IP 中。

设备配置如下。

第 1 步：Router1 的 VRRP 负载均衡配置。

```
Router1(config)#interface FastEthernet 0/0
Router1(config-if)#ip address 192.168.1.1 255.255.255.0
Router1(config-if)#vrrp 10    ip 192.168.1.253
Router1(config-if)#vrrp 10    priority 120
Router1(config-if)#vrrp 20    ip 192.168.1.254
Router1(config-if)#end
```

第 2 步：Router2 的 VRRP 负载均衡配置。

```
Router2(config)#interface FastEthernet 0/0
Router2(config-if)#ip address 192.168.1.2 255.255.255.0
Router2(config-if)#vrrp 10    ip 192.168.1.253
Router2(config-if)#vrrp 20    ip 192.168.1.254
Router1(config-if)#vrrp 20    priority 120
Router2(config-if)#end
```

第 3 步：在 Router1 上查看 VRRP 负载均衡状态。

```
Router1#show vrrp brief
Interface          Grp  Pri  Time  Own  Pre  State    Master addr     Group addr
FastEthernet 0/0   10   120  3     -    P    Master   192.168.1.1     192.168.1.253
FastEthernet 0/0   20   100  3     -    P    Backup   192.168.1.2     192.168.1.254
```

从 Router1 的显示结果可以看到，Router1 在 Group10 中为主路由器，在 Group20 中为备份路由器。

第 4 步：在 Router2 上查看 VRRP 负载均衡状态。

```
Router2#show vrrp brief
Interface          Grp  Pri  Time  Own  Pre  State    Master addr     Group addr
FastEthernet 0/0   10   100  3     -    P    Backup   192.168.1.1     192.168.1.253
FastEthernet 0/0   20   120  3     -    P    Master   192.168.1.2     192.168.1.254
```

从 Router2 的显示结果可以看到，Router2 在 Group20 中为主路由器，在 Group10 中为备份路由器。

7．VRRP 的监控与维护

（1）显示 VRRP 运行状态。

其中，*brief* 表示显示 VRRP 的概要信息，如果不指定参数，则显示所有 VRRP 组的状态信息；*group* 表示显示特定 VRRP 组的状态信息。

（2）显示指定接口上的 VRRP 信息。

指定的端口类型及编号

ruijie#show vrrp interface *type number [brief]*

交换机名称为ruijie
处于特权模式　　*显示VRRP概况*

其中，*type* 为接口的类型；*number* 为接口的编号；*brief* 为可选参数，若使用将显示概况。

3.3　扩展知识

1. 配置 VRRP 接口跟踪

在如图 3.9 所示的网络拓扑中，在 LAN1 中通过 Router1 和 Router2 实现局域网双出口连接，Router1 连接的链路带宽为 1000M，成为主链路；Router2 连接的链路带宽为 100M，成为备份链路。配置 Router1 成为主路由器，Router2 路由器成为备份路由器，正常情况下，数据流量通过 Router1 向外传输。

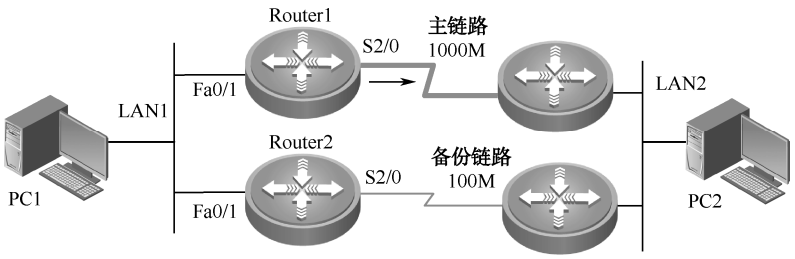


图 3.9　VRRP 接口跟踪

但是，当 Router1 的 S2/0 连接的链路出现故障时，Router1 仍作为主路由器从接口 Fa0/1 接收 LAN1 网络发送报文，然而，Router1 却无法对报文进行转发，使得网络无法正常通信。

为了解决这种问题，我们可以使用 VRRP 的接口跟踪机制。接口跟踪能够使 VRRP 根据路由器其他接口的状态自动调整该路由器 VRRP 优先级。当被跟踪接口不可用时，路由器的 VRRP 优先级将降低。接口跟踪能确保当主路由器的重要接口不可用时，该路由器不再是主路由器，从而使备份路由器有机会成为新的主路由器。

在如图 3.10 所示的网络拓扑中，Router1 的 VRRP 对 S2/0 接口进行跟踪。如果接口 S2/0 的链路出现故障，路由器会自动降低 VRRP 的优先级。这时如果 Router2 具有更高的优先级，Router2 将承担主路由器的角色。

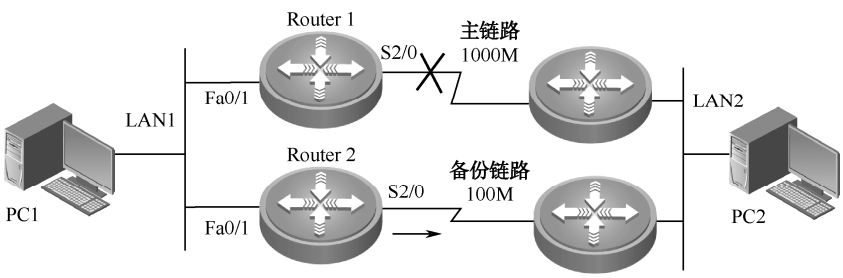


图 3.10　VRRP 路由器切换

在接口配置模式下，使用如下命令配置接口跟踪：

创建的VRRP组 *优先级减少值*

ruijie(config-if)# vrrp *group* **track** *interface-type number [priority]*

交换机名称为ruijie 处于接口配置模式 *监视的接口类型及接口号*

其中，*group* 为 VRRP 组号；*interface-type* 为被监视的接口类型；*number* 为被监视的接口号；*priority* 为被监视的接口状态改变时，其 VRRP 优先级的降低数值，默认情况下，降低数值为 10。当被跟踪接口恢复后，优先级也将恢复到原先的优先级值。

使用该命令的 no 选项，可以关闭该功能。如 ruijie(config-if)#no vrrp 10 track S 2/0。

提示：*在配置优先级的减少值时，需要保证降低后的优先级要小于现有备份路由器的优先级，以便让备份路由器接替主路由器的角色。*

例如：配置图 3.10 中 Router1 的接口跟踪功能，实现当 Serial 2/0 宕掉时，VRRP10 的优先级减少 30，切换到另一线路工作。

Router1 的主要配置代码如下。

```
Router1(config)#interface Serial 2/0
Router1(config-if)#ip address 200.1.1.1    255.255.255.0
Router1(config-if)#exit
Router1(config)#interface FastEthernet 0/1
Router1(config-if)#ip address 192.168.1.1 255.255.255.0
Router1(config-if)#vrrp 10    ip 10.1.1.253
Router1(config-if)#vrrp 10    priority 120
Router1(config-if)#vrrp 10    track Serial 2/0    30
Router1(config-if)#end
```

在 Router1 的配置中，为其配置了比默认优先级（100）高的优先级 120，并且被跟踪接口为 S2/0，当 S2/0 接口不可用时，减少的优先级为 30，即优先级降低到 90，这样可以保证具有更高优先级的 Router2（100）接替主路由器的角色。

使用 show vrrp 命令可以查看 VRRP 接口跟踪信息。

以下实例为 Router1 的 S2/0 链路不可用后 Router1 的 VRRP 状态。可以看到 Router1 的优先级已经降低为 90，并且状态为 Backup。

显示被跟踪接口失效后的 Router 1 的状态如下。

```
Router1#show vrrp brief
Interface          Grp   Pri   Time   Own   Pre   State     Master addr      Group addr
Ethernet 0/1       10    90    3      -     P     Backup    192.168.1.2      10.1.1.253
```

2．配置 VRRP 抢占模式

在 VRRP 运行过程中，主路由器定期地发送 VRRP 通告信息，备份路由器将侦听主路由器的通告信息。当备份路由器在主路由器失效间隔（master-down-interval）内没有接收到主路由器的通告消息时，它将认为主路由器失效，并接替后者的角色。

所谓的 VRRP 抢占（Preempt）模式，是指当原先的主路由器从故障中恢复并接入到网络中后，它将夺回原先属于自己的角色（主路由器）。如果不使用抢占模式，它从故障恢复后将保持备份路由器的状态。

在 VRRP 运行过程中，我们推荐启用抢占模式，这样可以使主链路恢复故障后，数据仍然通过主

链路传输。例如在使用一条高带宽链路和低带宽的备份链路的场景中，结合 VRRP 接口跟踪功能，可以使高带宽链路故障恢复后，仍然作为转发数据的主要链路，而不是使用低带宽的备份链路。

在接口配置模式下，使用如下命令配置 VRRP 抢占模式：

其中，*group* 为 VRRP 组号；*seconds* 为抢占的延迟时间，即发送通告报文前等待的时间，单位为秒，取值范围为 1～255。

默认情况下，VRRP 组工作在抢占模式，并且如果不配置延迟时间，那么默认值为 0 秒，即当路由器从故障中恢复后，立即进行抢占操作。

使用该命令的 no 选项将禁止 VRRP 功能，如 ruijie(config-if)#no vrrp 10 preempt。

例如：在如图 3.10 所示网络拓扑中，Router1 配置接口跟踪和抢占模式。当 Router1 的 S2/0 接口链路失效后，Router1 的优先级将降低到 90，并成为备份路由器，Router2 成为主路由器。当 Router1 的 S2/0 接口链路恢复后，Router1 的优先级将改变回原先的 120，并且由于启用了抢占模式，它将重新接替主路由器的角色。

Router1 主要配置代码如下：

```
Router1(config)#interface Serial 2/0
Router1(config-if)#ip address 200.1.1.1    255.255.255.0
Router1(config-if)#exit
Router1(config)#interface FastEthernet 0/1
Router1(config-if)#ip address 192.168.1.1 255.255.255.0
Router1(config-if)#vrrp 10    ip 10.1.1.253
Router1(config-if)#vrrp 10    priority 120
Router1(config-if)#vrrp 10    track Serial 2/0 30
Router1(config-if)#vrrp 10    preempt
Router1(config-if)#end
```

3. 配置 VRRP 定时器

在介绍 VRRP 定时器中我们已经了解到，VRRP 路由器使用通告报文来进行选举与状态监测。当选举结束后，主路由器将定期地发送通告报文，备份路由器将进行监听。我们可以在接口配置模式下，使用如下命令修改 VRRP 通告报文的发送间隔（adver-interval）：

其中，*group-number* 为 VRRP 组号；*advertise-interval* 为通告报文的发送间隔，单位为秒，取值范围为 1～255，默认为 1 秒。

使用该命令的 no 选项，恢复系统默认设置，如 ruijie(config-if)#no vrrp 10 timers advertise。

配置通告间隔时需要注意，较小的通告间隔将会造成一定的带宽和系统资源的消耗，尤其是当路

由器加入了多个 VRRP 组时，但是较小的通告间隔将提供更快的故障检测和切换；较大的通告间隔会节省带宽和系统资源，但不能提供最快的故障检测和切换。

在网络链路质量较差的环境中，为了使路由器在收到正常的通告报文前就进行状态切换，我们可以调高定时器的值，这将有助于提高网络的稳定性。在正常的网络环境中，推荐使用默认的定时器值。

对于 VRRP 中的另外一个定时器，主路由器失效间隔（master-down-interval），我们不能通过命令来对其进行配置，主路由器失效间隔是通过通告报文的时间间隔来进行计算的，它的值为通告间隔的 3 倍。

例如：配置 VRRP 通告报文间隔时间为 2 秒。

```
Router1(config)#interface FastEthernet 0/1
Router1(config-if)#ip address 192.168.1.1 255.255.255.0
Router1(config-if)#vrrp 10    ip 192.168.1.1
Router1(config-if)#vrrp 10    timers advertise    2
Router1(config-if)#end
```

3.4 演示实例

1. 背景描述

某集团公司的分公司与总公司之间网络通过专线连接，为了确保分公司与总公司之间可靠通信，采用两条链路相互备份。分公司与总公司网络运行 OSPF 协议。在分公司网络中采用 VRRP 技术实现网关冗余备份及负载均衡，如图 3.11 所示。

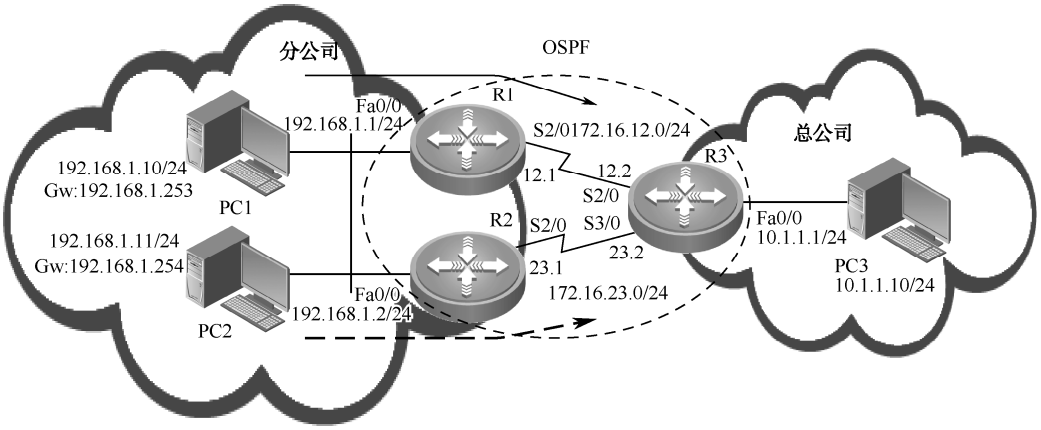


图 3.11　VRRP 操作演示

2. 操作步骤

第 1 步：在路由器上配置 IP 地址。

```
R1(config)#interface serial 2/0
R1(config-if)#ip address 172.16.12.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface FastEthernet 0/0
R1(config-if)#ip address 192.168.1.1 255.255.255.0
R1(config-if)#exit
```

```
R2(config)#interface serial 2/0
R2(config-if)#ip address 172.16.23.1 255.255.255.0
R2(config-if)#exit
R2(config)#interface FastEthernet 0/0
R2(config-if)#ip address 192.168.1.2 255.255.255.0
R2(config-if)#exit

R3(config)#interface serial 2/0
R3(config-if)#ip address 172.16.12.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface serial 3/0
R3(config-if)#ip address 172.16.23.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface FastEthernet 0/0
R3(config-if)#ip address 10.1.1.1 255.255.255.0
R3(config-if)#exit
```

第 2 步：配置 OSPF。

```
R1(config)#router ospf 1
R1(config-router)#network 192.168.1.0 0.0.0.255 area 0
R1(config-router)#network 172.16.12.0 0.0.0.255 area 0
R1(config-router)#exit

R2(config)#router ospf 1
R2(config-router)#network 192.168.1.0 0.0.0.255 area 0
R2(config-router)#network 172.16.23.0 0.0.0.255 area 0
R2(config-router)#exit

R3(config)#router ospf 1
R3(config-router)#network 10.1.1.0 0.0.0.255 area 0
R3(config-router)#network 172.16.12.0 0.0.0.255 area 0
R3(config-router)#network 172.16.23.0 0.0.0.255 area 0
R3(config-router)#exit
```

第 3 步：配置 VRRP。

在路由器 R1 上创建 VRRP10 虚拟组，虚拟组 IP 地址为 192.168.1.253，优先级为 120。当路由器 S2/0 接口状态为 Down 时，路由器优先级降低 30。创建 VRRP20 虚拟组，虚拟组 IP 地址为 192.168.1.254，优先级为默认值 100。

```
R1(config)#interface FastEthernet 0/0
R1(config-if)#vrrp 10 ip 192.168.1.253
R1(config-if)#vrrp 10 priority 120
R1(config-if)#vrrp 10 track serial 2/0 30
R1(config-if)#vrrp 10 ip 192.168.1.254
```

在路由器 R2 上创建 VRRP10 虚拟组，虚拟组 IP 地址为 192.168.1.253，优先级为默认值 100；创

建 VRRP20 虚拟组，虚拟组 IP 地址为 192.168.1.254，优先级为默认值 120。当路由器 S2/0 接口状态为 Down 时，路由器优先级降低 30。

```
R2(config)#interface FastEthernet 0/0
R2(config-if)#vrrp 10 ip 192.168.1.253
R2(config-if)#vrrp 20 ip 192.168.1.254
R2(config-if)#vrrp 20 priority 120
R2(config-if)#vrrp 20 track serial 2/0 30
```

第 4 步：验证测试。

使用 show vrrp brief 命令来验证配置是否正确。

```
R1#show vrrp brief
Interface          Grp   Pri   Time   Own   Pre   State      Master addr     Group addr
FastEthernet 0/0   10    120   -      -     P     Master     192.168.1.1     192.168.1.253
FastEthernet 0/0   20    100   -      -     P     Backup     192.168.1.2     192.168.1.254
```

从上述显示结果看出，R1 路由器在 VRRP10 组中优先级为 120，状态为 Master 路由器，在 VRRP20 组中优先级为 100，状态为 Backup 路由器。

```
R2#show vrrp brief
Interface          Grp   Pri   Time   Own   Pre   State      Master addr     Group addr
FastEthernet 0/0   10    100   -      -     P     Backup     192.168.1.1     192.168.1.253
FastEthernet 0/0   20    120   -      -     P     Master     192.168.1.2     192.168.1.254
```

从上述显示结果看出，R2 路由器在 VRRP10 组中优先级为 100，状态为 Backup 路由器，在 VRRP20 组中优先级为 120，状态为 Master 路由器。

在 R3 路由器的 S2/0 接口上使用 shutdown 命令关闭该接口，命令如下。

```
R3(config)#interface serial 2/0
R3(config-if)#shutdown
R3(config-if)#
```

R1 路由器的 serial 2/0 接口的状态也变为 Down。此时再使用 show vrrp brief 命令查看接口的 VRRP 信息如下。

```
R1#show vrrp brief
Interface          Grp   Pri   Time   Own   Pre   State      Master addr     Group addr
FastEthernet 0/0   10    90    -      -     P     Backup     192.168.1.2     192.168.1.253
FastEthernet 0/0   20    100   -      -     P     Backup     192.168.1.2     192.168.1.254
```

从上述显示结果看出，当监控接口状态变为 Down 时，R1 路由器在 VRRP10 组中优先级从 120 降至 90，状态由 Master 路由器变为 Backup 路由器，在 VRRP20 组中优先级仍为 100，状态仍为 Backup 路由器。

```
R2#show vrrp brief
Interface          Grp   Pri   Time   Own   Pre   State      Master addr     Group addr
FastEthernet 0/0   10    100   -      -     P     Master     192.168.1.2     192.168.1.253
FastEthernet 0/0   20    120   -      -     P     Master     192.168.1.2     192.168.1.254
```

从上述显示结果看出，路由器 R2 优先级仍为 100，但由于路由器 R1 的优先级降为 90，使得 R2 路由器在 VRRP10 中状态从 Backup 路由器变为 Master 路由器。在 VRRP20 组的状态不变。

3.5　训练任务

1．背景描述

某公司局域网使用双核心交换机与路由器连接，实现局域网主机访问 Internet。为了保证主机网关冗余备份及负载均衡，在两台核心交换机上启动 VRRP 功能，如图 3.12 所示。

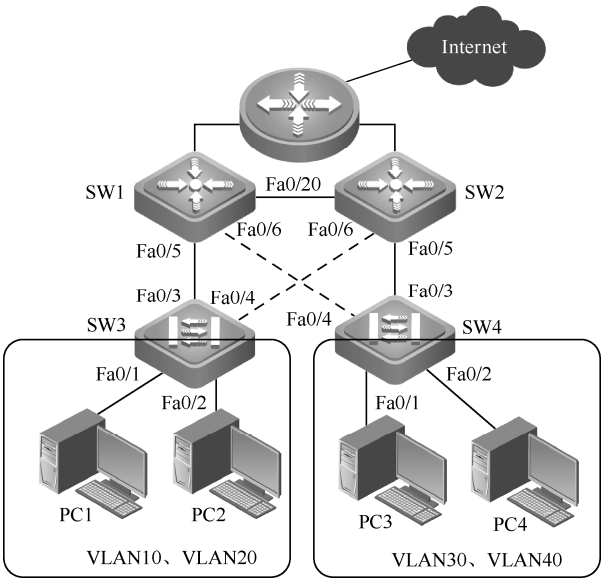


图 3.12　VRRP 操作演示

通过对交换机 SW1 和 SW2 配置，使得 SW1 对于 VLAN10、VLAN20 的优先级为 255，对于 VLAN30、VLAN40 的优先级为 100，成为 VLAN10、VLAN20 的 Master；SW2 对于 VLAN10、VLAN20 的优先级为 100，对于 VLAN30、VLAN40 的优先级为 255，成为 VLAN30、VLAN40 的 Master。

2．操作提示

第 1 步：在交换机 SW1 上创建 VRRP10、VRRP20、VRRP30 及 VRRP40 虚拟组，并将 VRRP10、VRRP20 的虚拟组 IP 地址设置为相应的 SVI 地址，优先级为 255。同时将 VRRP10、VRRP20 虚拟组设置为抢占模式。参考配置代码如下：

```
SW1(config)#interface VLAN 10
SW1(config-if)#ip address 192.168.10.1 255.255.255.0
SW1(config-if)#vrrp 10 ip 192.168.10.1
SW1(config-if)#vrrp 10 preempt
SW1(config)#interface VLAN 20
SW1(config-if)#ip address 192.168.20.1 255.255.255.0
SW1(config-if)#vrrp 20 ip 192.168.20.1
SW1(config-if)#vrrp 20 preempt
SW1(config)#interface VLAN 30
SW1(config-if)#ip address 192.168.30.1 255.255.255.0
SW1(config-if)#vrrp 30 ip 192.168.30.2
```

```
SW1(config)#interface VLAN 40
SW1(config-if)# ip address 192.168.40.1 255.255.255.0
SW1(config-if)#vrrp 40 ip 192.168.40.2
```

第 2 步：在交换机 SW2 上创建 VRRP10、VRRP20、VRRP30 及 VRRP40 虚拟组，并将 VRRP30、VRRP40 的虚拟组 IP 地址设置为相应的 SVI 地址，优先级为 255。同时将 VRRP30、VRRP40 虚拟组设置为抢占模式。参考配置代码如下：

```
SW2(config)#interface VLAN 10
SW2(config-if)#ip address 192.168.10.2 255.255.255.0
SW2(config-if)#vrrp 10 ip 192.168.10.1
SW2(config)#interface VLAN 20
SW2(config-if)#ip address 192.168.20.2 255.255.255.0
SW2(config-if)#vrrp 20 ip 192.168.20.1
SW2(config)#interface VLAN 30
SW2(config-if)#ip address 192.168.30.2 255.255.255.0
SW2(config-if)#vrrp 30 ip 192.168.30.2
SW2(config-if)#vrrp 30 preempt
SW2(config)#interface VLAN 40
SW2(config-if)#ip address 192.168.40.2 255.255.255.0
SW2(config-if)#vrrp 40 ip 192.168.40.2
SW2(config-if)#vrrp 40 preempt
```

练习题

1．选择题

（1）交换机启动 VRRP 协议后，若主交换机需要关闭 VRRP 协议，则其发布通告的优先级将（　　）。

A．变为 0　　B．变为 1　　C．变为 255　　D．不变

（2）在交换机上开启 VRRP，若备用交换机从 Backup 状态转换为 Master 状态，则最不可能的原因是（　　）。

A．Master 与 Backup 之间链路 Down　　B．Master 优先级低于或等于 Backup 优先级

C．Master 优先级变为 0　　D．Master_down_time 计时器超时

（3）在交换机上开启 VRRP 功能，若网络中时延较大，则（　　）能防止主备频繁倒换。

A．修改 VRRP 发送通告报文的时间间隔　　B．修改 VRRP 的 Track 值

C．设置 VRRP 实体间的认证　　D．开启备份组抢占方式

2．简答题

（1）描述 VRRP 作用及应用场合。

（2）VRRP 报文的组播地址是多少？协议号是多少？

（3）说明如何实现 VRRP 负载均衡。

（4）说明如何跟踪链路状态，调整优先级。

（5）在 MSTP 与 VRRP 网络中，如何合理地设置 Root 交换机及 Master 交换机？

项目 4　动态主机配置协议 DHCP

当在网络中主机数量比较多或主机需临时接入网络（如无线网络构成的会场、机场等环境）的情况下，如果主机网络参数手工配置的话，网络管理员的工作量会很大，同时也容易产生错误。在这种情况下，宜采用 DHCP 网络参数自动配置。

通过本章对 DHCP 协议工作原理进行简单介绍，同时描述 DHCP 配置的相关命令的使用方法。

知识点、技能点

1. 了解 DHCP 工作原理;
2. 掌握 DHCP 特性及配置技能。

4.1　问题提出

某单位有 500 台计算机，通过交换机、路由器搭建成局域网，为了方便用户使用，要求用户端网络参数自动配置，你作为网络管理员应如何实现呢？

4.2　相关知识

1. DHCP 概述

DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）是一种在网络中常用的动态配置网络参数的技术，代替网络管理员手工配置及维护 IP 地址的工作。

DHCP 使用 UDP 传输协议，使用 67、68 接口号，从 DHCP 客户端到达 DHCP 服务器的报文使用目的接口号 67，从 DHCP 服务器到达 DHCP 客户端报文使用源接口号 68。

DHCP 的工作过程如下。

（1）发现阶段。发现阶段是 DHCP 客户端寻址 DHCP 服务器的阶段。在这期间，由于 DHCP 客户端还没有可用的 IP 地址，那么 DHCP 客户端使用目的 IP 地址 255.255.255.255 和目的 MAC 地址 FFFF-FFFF_FFFF 以广播方式发送 DHCPDISCOVER 发现报文。网络中的每一台主机（包括所有客户端主机及 DHCP 服务器）都会收到这个发现报文，如图 4.1 所示。

（2）提供阶段。提供阶段是 DHCP 服务器提供 IP 地址的阶段。DHCP 服务器收到 DHCPDISCOVER 报文后，从 DHCP 服务器中尚未分配的 IP 地址中选出一个 IP 地址分配给 DHCP 客户端，并通过向 DHCP 客户端发送一个包含待分配的 IP 地址和其他网络参数的 DHCPOFFER 提供报文作为对客户端的响应，如图 4.2 所示。

（3）选择阶段。选择阶段是 DHCP 客户端对 DHCP 服务器提供的 IP 地址进行选择的阶段。如果网络中有多台 DHCP 服务器同时提供 DHCP 服务，DHCP 客户端只接收第一个收到的 DHCPOFFER 报文，然后，DHCP 客户端以广播方式发送 DHCPREQUEST 请求报文作为对 DHCP 服务器的响应，如图 4.3 所示。

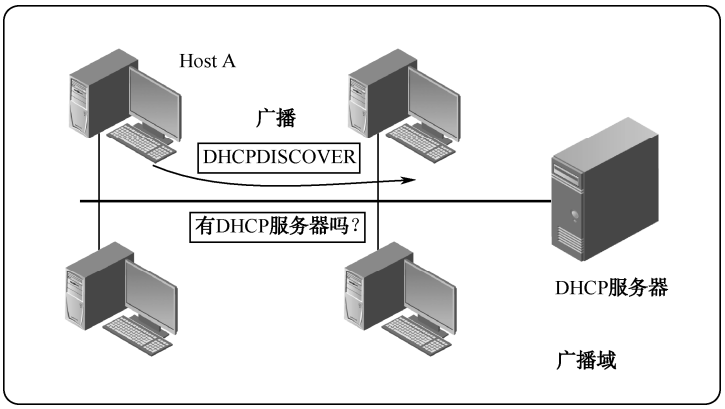


图 4.1　DHCP 发送发现报文

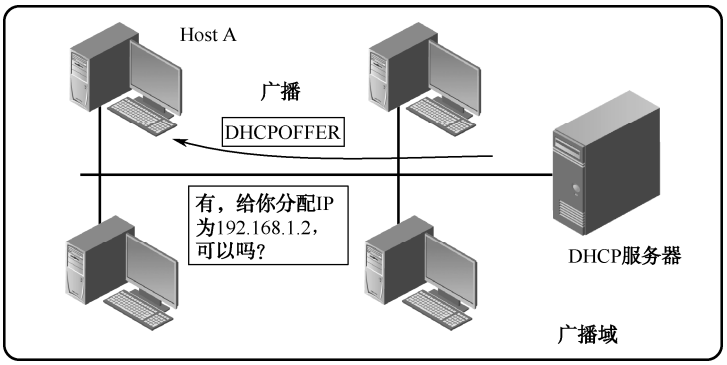


图 4.2　DHCP 发送提供报文

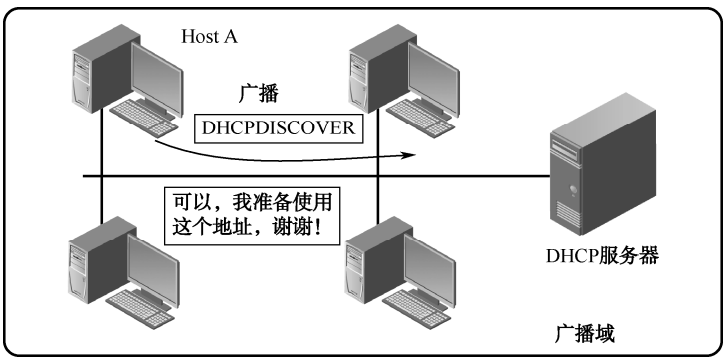


图 4.3　DHCP 发送选择报文

（4）确认阶段。确认阶段是 DHCP 服务器向客户端确认所提供 IP 地址可以使用的阶段。当 DHCP 服务器收到 DHCP 客户端的 DHCPREQUEST 请求报文后，向客户端发送一个包含所提供 IP 地址及其他网络参数的 DHCPACK 确认报文，表明客户端可以使用所提供的网络参数。DHCP 客户端收到 DHCPACK 确认报文后，开始使用 IP 地址及其他网络参数进行网络配置，如图 4.4 所示。

2．配置 DHCP 服务器

DHCP 服务器既可以在 Windows 或 Linux 等操作系统上构建，也可以在网络设备（如路由器、三层交换机等）上构建。以下只介绍在网络设备上构建 DHCP 的相关命令。

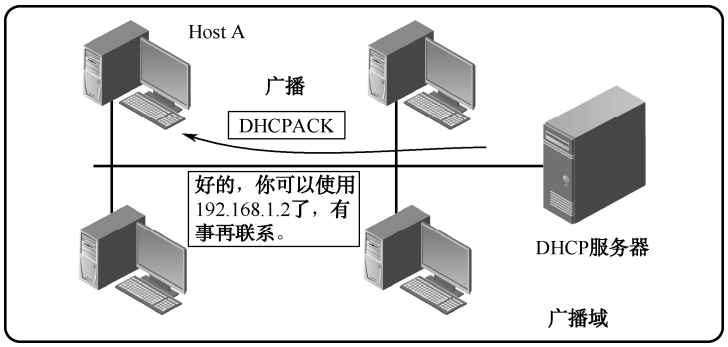


图 4.4　DHCP 发送确认报文

（1）启用 DHCP 服务器。默认情况下，路由器或三层交换机未启用 DHCP 服务器和中继代理特性。若使用 DHCP 服务器，必须启动 DHCP 服务，具体配置的命令格式如下：

启动DHCP服务及中继代理功能

ruijie(config)# service dhcp

路由器名称为ruijie
处于全局配置模式

提示：*使用该命令的 no 选项可以关闭 DHCP 服务器和中继代理特性，如 ruijie(config)#no service dhcp。*

例如：启动 DHCP 服务器和中继代理特性。

```
ruijie(config)#service dhcp
```

（2）定义 DHCP 排除地址范围。默认情况下，DHCP 服务器会试图将在地址池中定义的所有子网地址分配给 DHCP 客户端。如果想保留一些地址不分配，比如已经分配给服务器或者设备了，则必须明确定义这些地址是不允许分配给客户端的。

定义 DHCP 排除地址范围的命令格式如下：

排除地址范围的起始地址

ruijie(config)#ip dhcp excluded-address *low-ip-address [high-ip-address]*

路由器名称为ruijie
处于全局配置模式

排除地址范围的结束地址

其中，*low-ip-address* 为排除 IP 地址，或者排除 IP 地址范围的起始 IP 地址；*high-ip-address* 为排除地址范围的结束 IP 地址。默认情况下，DHCP 服务器分配整个地址池的 IP 地址。

提示：*该命令的 no 选项可以取消这种定义，如 ruijie(config)#no ip dhcp excluded-address 192.168.1.10 192.168.1.20。*

例如：DHCP 服务器将不会分配 192.168.12.100～19.2168.12.150 范围内的 IP 地址。

```
ruijie(config)#ip dhcp excluded-address 192.168.12.100 192.168.12.150
```

（3）配置 DHCP 地址池。DHCP 的地址分配及给客户端传送的 DHCP 各项参数，都需要在 DHCP 地址池中进行定义。如果没有配置 DHCP 地址池，即使启动了 DHCP 服务器，也不能对客户端进行地址分配；但是如果启用了 DHCP 服务器，不管是否配置了 DHCP 地址池，DHCP 中继代理总是起作用的。配置 DHCP 地址池的命令如下。

第 1 步：创建 DHCP 地址池。首先，要定义一个 DHCP 地址池名并进入 DHCP 地址池的配置模

式，然后才能在地址池配置模式下，进行相关参数配置。创建 DHCP 地址池的命令格式如下：

其中，*pool-name* 为地址池名称，可以由字符和正整数组成。如 mypool，或者 1。

默认情况下，没有定义 DHCP 地址池。使用该命令的 no 选项可以删除 DHCP 地址池，如 ruijie(config)#no ip dhcp pool mypool。

例如：定义一个名为 mypool0 的 DHCP 地址池。

```
ruijie(config)#ip dhcp pool mypool0
ruijie(dhcp-config)#
```

例如：删除已定义的 mypool0 地址池。

```
ruijie(config)#no ip dhcp pool mypool0
```

第 2 步：定义地址范围。定义地址池中 IP 地址使用范围，为 DHCP 服务器提供了一个可分配给客户端的地址空间。除非有地址排除配置，否则地址池中的所有地址都有可能分配给客户端。DHCP 在分配地址池中的地址是按顺序进行的，如果该地址已经在 DHCP 绑定表中或者检测到该地址已经在该网段中存在，就检查下一个地址，直到分配一个有效的地址。

定义地址范围的命令格式如下：

其中，*network-number* 为 DHCP 地址池的 IP 地址网络号。*net-mask* 为 DHCP 地址池的 IP 地址网络掩码。如果没有定义掩码，默认为自然网络掩码。

使用该命令的 no 选项可以删除定义，如 ruijie(dhcp-config)#no network。默认情况下，没有定义网络号和网络掩码。

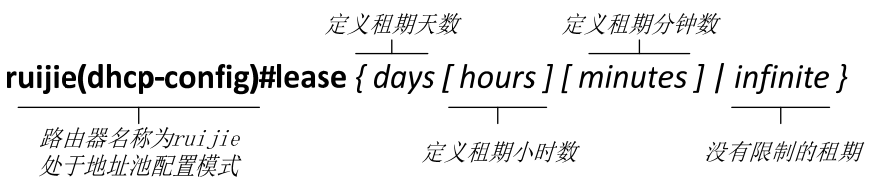
例如：定义 DHCP 地址池的网络号为 192.168.12.0，掩码为 255.255.255.240。

```
ruijie(dhcp-config)#network 192.168.12.0 255.255.255.240
```

例如：删除已定义的 DHCP 地址池中的网络号和子网掩码。

```
ruijie(dhcp-config)#no network
```

第 3 步：定义地址租期。DHCP 服务器分配给客户端的 IP 地址是有时间限制的，这个时间称为地址租期。当租期快到时，DHCP 客户端会发送续租请求。DHCP 服务器一般会允许续租，续租的地址保持不变。设置租期的命令格式如下：



其中，*days* 为以天为单位定义的租期时间；*hours*（可选）为以小时为单位定义的租期时间，定义小时数前必须定义天数；*minutes*（可选）为以分钟为单位定义的租期时间，定义分钟前必须定义天数和小时数；*infinite* 表示没有限制的租期。

默认情况下，租期为 1 天。使用该命令的 no 选项恢复默认值，如 ruijie(dhcp-config)#no lease。

例如：将 DHCP 租期设为 1 小时。

```
ruijie(dhcp-config)#lease 0 1
```

例如：将 DHCP 租期设为 1 分钟。

```
ruijie(dhcp-config)#lease 0 0 1
```

例如：将 DHCP 租期恢复为默认值。

```
ruijie(dhcp-config)#no lease
```

第 4 步：设置默认网关。一般情况下，DHCP 客户端需要从 DHCP 服务器获得默认网关信息。DHCP 服务器至少需要为客户端指定一个网关 IP 地址，该地址必须与分配给客户的地址在同一个网段。如果配置了多个地址，写在前面的网关具有更高的优先权。设置默认网关的命令格式如下：

ruijie(dhcp-config)#default-router *ip-address [ip-address2…ip-address8]*

路由器名称为ruijie
处于地址池配置模式

设置默认网关IP地址

其中，*ip-address* 为设置的默认网关主 IP 地址；*ip-address2…ip-address8*（可选）为设置的默认网关次 IP 地址，最多可以配置 7 个次网关。

使用该命令的 no 选项可以删除默认网关 IP 地址的定义，如 ruijie(dhcp-config)#no default-router。默认情况下，没有定义默认网关 IP 地址。

例如：定义 192.168.12.1 为默认网关 IP 地址。

```
ruijie(dhcp-config)#default-router 192.168.12.1
```

例如：删除已设置的 DHCP 默认网关 IP 地址。

```
ruijie(dhcp-config)#no default-router
```

第 5 步：设置客户端后缀域名。DHCP 客户端获得了指定后缀域名后，当访问具有同样后缀域名的主机时，直接通过主机名就可以了。

设置客户端后缀域名的命令格式如下：

定义的客户端后缀域名

ruijie(dhcp-config)#domain-name *domain-name*

路由器名称为ruijie
处于地址池配置模式

其中，*domain-name* 为设置的 DHCP 客户端的后缀域名字符串。

使用该命令的 no 选项可以删除后缀域名，如 ruijie(dhcp-config)#no domain-name。默认情况下，没有后缀域名。

例如：给 DHCP 客户端定义后缀域名为 i-net.com.cn。

```
ruijie(dhcp-config)#domain-name i-net.com.cn
```

例如：删除已设置的客户端后缀域名。

```
ruijie(dhcp-config)#no domain-name
```

第 6 步：指定 DNS 服务器。定义多个 DNS 服务器时，写在前面的优先权高，DHCP 客户端只有

与排在前头的 DNS 服务器通信失败时，才会选择下一个 DNS 服务器。

指定 DNS 服务器命的令格式如下：

DNS服务器IP地址

ruijie(dhcp-config)#dns-server *{ ip-address [ip-address2…ip-address8] }*

路由器名称为ruijie
处于地址池配置模式

其中，*ip-address* 为指定的 DNS 服务器主 IP 地址；*ip-address2…ip-address8*（可选）为指定的 DNS 服务器的次 IP 地址，最多可以配置 7 个 DNS 服务器。

使用该命令的 no 选项可以删除 DNS 服务器的定义，如 ruijie(dhcp-config)#no dns-server。默认情况下，没有定义 DNS 服务器。

例如：为 DHCP 客户端指定 DNS 服务器为 192.168.12.3。

```
ruijie(dhcp-config)#dns-server 192.168.12.3
```

例如：删除已定义的 DNS 服务器 IP 地址。

```
ruijie(dhcp-config)#no dns-server
```

第 7 步：指定 WINS 服务器。定义多个 WINS 服务器时，写在前面的优先权高，DHCP 客户端只有与排在前头的 WINS 服务器通信失败时，才会选择下一个 WINS 服务器。

指定 WINS 服务器的命令格式如下：

WINS服务器IP地址

ruijie(dhcp-config)#netbios-name-server *ip-address [ip-address2…ip-address8]*

路由器名称为ruijie
处于地址池配置模式

其中，*ip-address* 为定义的 WINS 服务器 IP 地址；*ip-address2…ip-address8* 为（可选）定义的 WINS 服务器次 IP 地址，最多可以配置 7 个 WINS 服务器。

使用该命令的 no 选项可以删除 WINS 服务器，如 ruijie(dhcp-config)#no netbios-name-server。默认情况下，没有定义 WINS 服务器。

例如：为 DHCP 客户端指定 WINS 服务器 IP 地址为 192.168.12.3。

```
ruijie(dhcp-config)#netbios-name-server 192.168.12.3
```

例如：删除已定义的客户端 WINS 服务器。

```
ruijie(dhcp-config)#no netbios-name-server
```

第 8 步：指定客户端 NetBIOS 节点类型。微软 DHCP 客户端 NetBIOS 节点有如下 4 种类型。

① Broadcast，广播型节点，通过广播方式进行 NetBIOS 名字解析。

② Peer-to-peer，对等型节点，通过直接请求 WINS 服务器进行 NetBIOS 名字解析。

③ Mixed，混合型节点，先通过广播方式请求名字解析，后通过与 WINS 服务器连接进行名字解析。

④ Hybrid，复合型节点，首先直接请求 WINS 服务器进行 NetBIOS 名字解析，如果没有得到应答，就通过广播方式进行 NetBIOS 名字解析。

设置客户端 NetBIOS 节点类型的命令格式如下。

NetBIOS节点类型

ruijie(dhcp-config)#netbios-node-type *type*

路由器名称为ruijie
处于地址池配置模式

其中，*type* 为定义的 NetBIOS 节点类型。定义节点类型有数字和字符串两种方式，数字方式有：1，代表 b-node；2，代表 p-node；4，代表 m-node；8，代表 h-node。字符串方式有：b-node，广播型节点；p-node，对等型节点；m-node，混合型节点；h-node，复合型节点。

使用该命令的 no 选项可以删除 NetBIOS 节点类型配置，如 ruijie(dhcp-config)#no netbios-mode-type。默认情况下，没有定义 NetBIOS 节点类型。

例如：将微软 DHCP 客户端的 NetBIOS 节点设为复合型节点。

```
ruijie(dhcp-config)#netbios-node-type h-node
```

例如：删除已定义的 NetBIOS 节点类型。

```
ruijie(dhcp-config)#no netbios-node-type
```

例如：某公司网络中定义了一个地址池 net172，地址池中的地址范围为 172.16.1.0/24，默认网关为 172.16.1.254，域名为 rg.com，域名服务器为 172.16.1.253，WINS 服务器为 172.16.1.252，NetBIOS 节点类型为复合型，地址租期为 30 天。该地址池中除了 172.16.1.2～172.16.1.100 地址外，其余地址均为可分配地址，如图 4.5 所示。

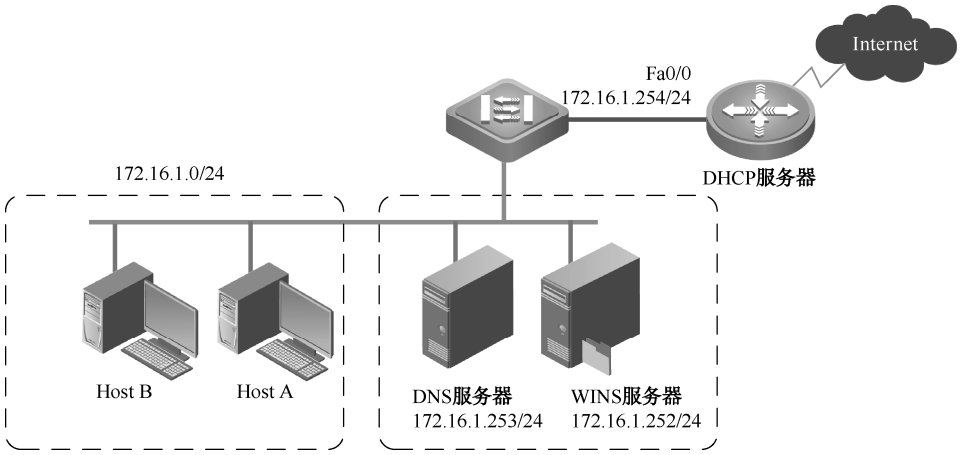


图 4.5　DHCP 服务器应用

具体配置如下：

```
ruijie(config)#ip dhcp excluded-address 172.16.1.2 172.16.1.100
ruijie(config)#ip dhcp excluded-address 172.16.1.252 172.16.1.254
ruijie(config)#ip dhcp pool net172
ruijie(dhcp-config)#network 172.16.1.0 255.255.255.0
ruijie(dhcp-config)#default-router 172.16.1.254
ruijie(dhcp-config)#domain-name rg.com
ruijie(dhcp-config)#dns-server 172.16.1.253
ruijie(dhcp-config)#netbios-name-server 172.16.1.252
ruijie(dhcp-config)#netbios-node-type h-node
```

```
ruijie(dhcp-config)#lease 30
ruijie(dhcp-config)#exit
ruijie(config)#
```

3. DHCP 中继代理工作原理

DHCP 客户端通过广播方式发送 DHCPDISCOVER 发现报文，发现 DHCP 服务器，实现 DHCP 服务功能。在大型网络中，往往划分了多个子网，每个子网都是一个广播域，DHCPDISCOVER 发现报文到达子网网关终止，如果 DHCP 客户端及服务器不在同一个广播域内，不能实现 DHCP 服务功能。若要在这种情况下，实现 DHCP 服务功能，则需要设置 DHCP 中继代理。

DHCP 中继代理可以在路由器或三层交换机上设置，在 DHCP 客户端及服务器间中转相关报文。DHCP 客户端与代理间仍采用广播方式，DHCP 服务器与代理间采用单播方式，如图 4.6 所示。

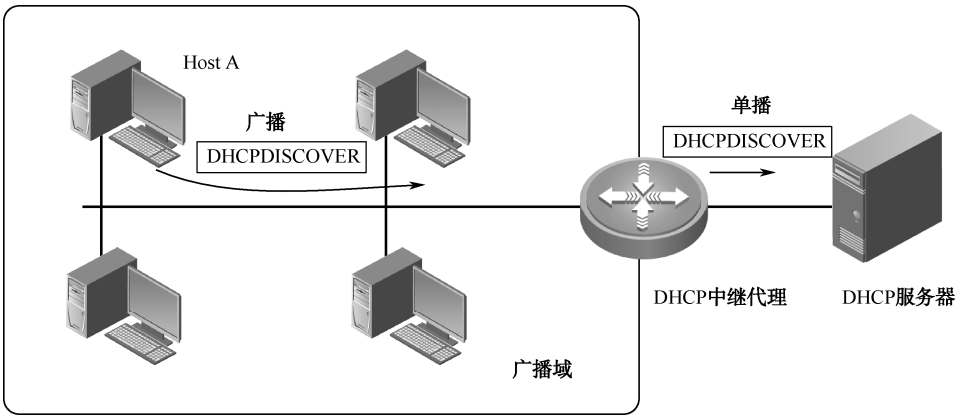


图 4.6　DHCP 中继代理

具体工作过程如下。

（1）发现阶段。客户端使用广播方式发送 DHCPDISCOVER 发现报文，DHCP 中继代理收到后，记录客户端子网地址信息，并以单播方式转发 DHCPDISCOVER 发现报文到达 DHCP 服务器。

（2）提供阶段。当 DHCP 服务器收到 DHCPDISCOVER 发现报文后，根据客户端子网地址信息分配相应子网的地址，并以单播方式发送 DHCPOFFER 提供报文给中继代理。DHCP 中继代理收到 DHCPOFFER 提供报文后，以广播方式转发 DHCPOFFER 提供报文到达客户端子网。

（3）选择阶段。当客户端收到 DHCPOFFER 提供报文后，仍以广播方式发送 DHCPREQUEST 选择报文。DHCP 中继代理收到后，以单播方式发送 DHCPREQUEST 选择报文到达 DHCP 服务器。

（4）确认阶段。当 DHCP 服务器收到 DHCPREQUEST 选择报文后，以单播方式发送 DHCPACK 确认报文给 DHCP 中继代理。DHCP 中继代理以广播方式将 DHCPACK 确认报文转发给客户端，客户端使用 DHCP 服务器分配的 IP 地址及其他网络参数配置网络。

从以上工作过程可以看出，DHCP 中继代理在 DHCP 客户端及服务器间起到中转服务的功能，从而实现 DHCP 服务器为不同子网提供 DHCP 网络服务的目的。

4. 配置 DHCP 中继代理

（1）启用 DHCP 服中继代理。默认情况下，路由器未启用 DHCP 中继代理。若要使用 DHCP 服务器，则必须启动 DHCP 服务，配置命令同启动 DHCP 服务一样，具体格式如下：

启动DHCP服务及中继代理功能

ruijie(config)# service dhcp

路由器名称为ruijie
处于全局配置模式

（2）指定 DHCP 服务器的 IP 地址。在 DHCP 中继代理上配置 DHCP 服务器的 IP 地址后，DHCP 代理把所收到主机的请求报文将转发给 DHCP 服务器，同时，收到的来自 DHCP 服务器的响应报文也会转发给主机。

DHCP 服务器地址可以全局配置，也可以在三层接口上配置，每种配置模式都可以配置多个服务器地址，最多可以配置 20 个服务器地址。若在某接口收到 DHCP 请求，则首先使用接口 DHCP 服务器；如果接口上面没有配置服务器地址，则使用全局配置的 DHCP 服务器。

在全局配置模式下指定 DHCP 服务器 IP 地址的命令格式如下:

DHCP服务器IP地址

ruijie(config)# IP helper-address *A.B.C.D*

路由器名称为ruijie
处于全局配置模式

其中，*A.B.C.D* 为 DHCP 服务器地址。使用该命令的 no 选项可以删除一个 DHCP 服务器 IP 地址，如 ruijie(config)#no ip helper-address 192.168.1.100。默认情况下，没有配置任何 DHCP 服务器 IP 地址。

在接口配置模式下指定 DHCP 服务器 IP 地址的命令格式如下：

DHCP服务器IP地址

ruijie(config-if)# IP helper-address *A.B.C.D*

路由器名称为ruijie
处于接口配置模式

其中，*A.B.C.D* 为 DHCP 服务器地址。使用该命令的 no 选项可以删除一个 DHCP 服务器 IP 地址，如 ruijie(config-if)#no ip helper-address 192.168.1.200。默认情况下，没有配置任何 DHCP 服务器 IP 地址。

例如：某公司网络使用 DHCP 服务器进行动态地址分配。DHCP 服务器与客户端分别位于不同子网，如图 4.7 所示。客户端子网地址为 172.16.1.0/24，DHCP 服务器的 IP 地址为 192.168.1.252/24，DNS 服务器的 IP 地址为 192.168.1.253/24,，WINS 服务器的 IP 地址为 192.168.1.254/24。客户端的 NetBIOS 节点类型为复合型节点，地址租期为 4 天。现需要配置 DHCP 中继代理，使得客户端自动获得 IP 地址等网络参数。

DHCP 中继代理配置如下：

```
ruijie(config)#service dhcp
ruijie (config)#interface fastether net 0/0
ruijie(config-if)#ip address 172.16.1.1 255.255.255.0
ruijie(config-if)#ip helper-address 192.168.1.252
ruijie(config-if)#end
ruijie#
```

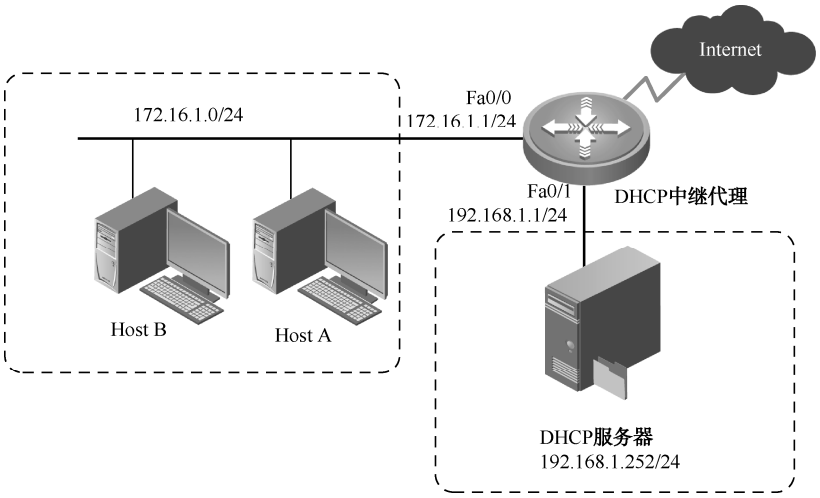


图 4.7　DHCP 中继代理应用

4.3　扩展知识

1. DHCP 服务器配置（Windows）

在 Windows Server 2008 系统中，默认情况下没有安装 DHCP 服务器，因此，为了提供 DHCP 服务，需要安装 DHCP 服务器。安装 DHCP 服务器的主机要使用固定 IP 地址（本例中使用 192.168.11.230/24），在安装 DHCP 服务器之前必须设置完毕。

安装与配置 DHCP 服务器的步骤如下。

（1）启动 Windows Server 2008 系统后，以 Administrator 身份登录。单击“开始”按钮，选择“管理工具”→“服务器管理”命令，进入“服务器管理器”窗口，单击“角色”选项，如图 4.8 所示。

图 4.8　添加服务器角色

（2）单击“添加角色”选项，然后单击“服务器角色”选项，进入“选择服务器角色”窗口，如图 4.9 所示。选中“DHCP 服务器”复选框，单击“下一步”按钮。

（3）进入“DHCP 服务器简介及注意事项”窗口，阅读后，单击“下一步”按钮。进入“选择网络连接绑定”窗口，选择 DHCP 服务器使用的 IP 地址，如图 4.10 所示，单击“下一步”按钮。

图 4.9　选择服务器角色

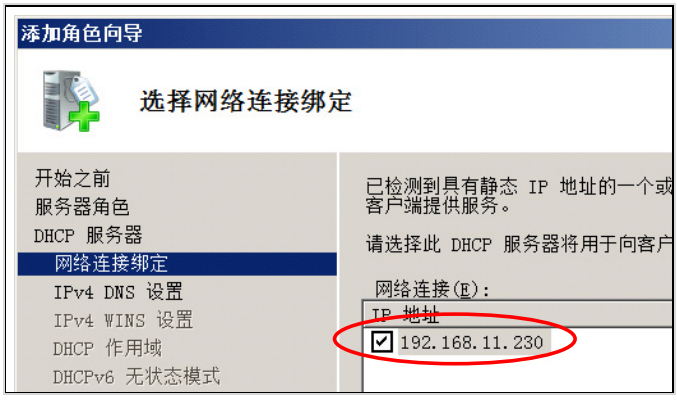


图 4.10　选择服务器 IP 地址

（4）进入“指定 IPv4 DNS 服务器设置”窗口，设置父域名称和 DNS 服务器 IP 地址，单击“下一步”按钮，如图 4.11 所示。

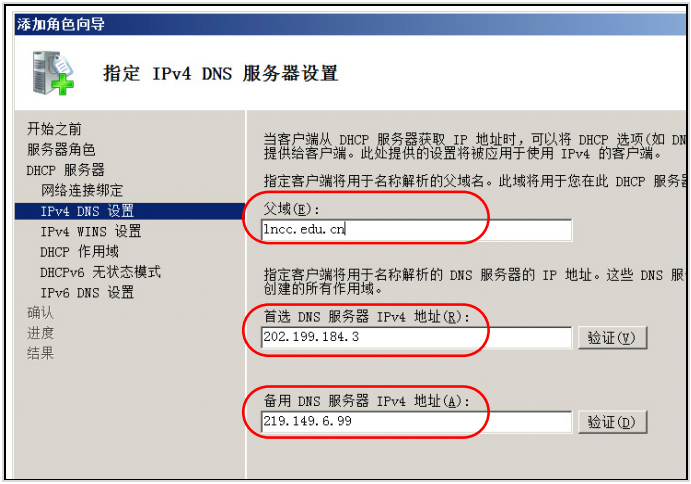


图 4.11　设置 DNS 服务器信息

（5）进入“指定 IPv4 WINS 服务器设置”窗口，此处我们不设置，直接单击“下一步”按钮，如图 4.12 所示。

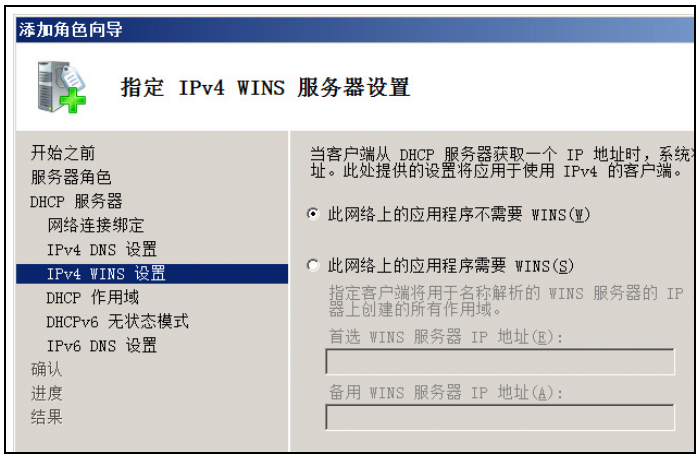


图 4.12　设置 WINS 服务器信息

（6）进入“添加作用域”窗口，设置起始 IP 地址、结束 IP 地址、子网掩码和默认网关的值，选中“激活此作用域”选项，然后单击“下一步”按钮，如图 4.13 所示。

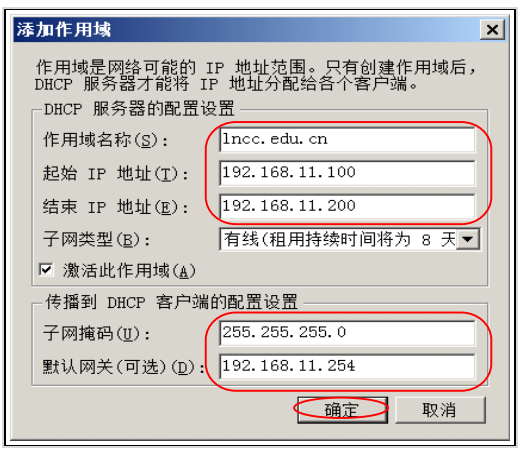


图 4.13　设置作用域信息

（7）此时，系统出现 IPv6 设置窗口，可以根据实际情况进行设定。在此我们不涉及 IPv6，因此直接单击两次“下一步”按钮，跳过对 IPv6 的设置。

（8）系统弹出“确认安装选择”窗口，阅读窗口列出设置的信息，确认无误后单击“安装”按钮，如图 4.14 所示。

（9）安装成功后，显示如图 4.15 所示信息，表示安装成功。单击“关闭”按钮。

2．客户端配置（Windows）

（1）在客户端主机上使用 Windows 7 操作系统。启动 Windows 7 操作系统后，单击“开始”，选择“控制面板”命令，在打开的窗口中单击“网络和共享中心”选项，再单击“网络适配器设置”选项。

（2）进入“网络适配器设置”窗口后，选择并右击“网络连接”图标，在弹出的快捷菜单中选择“属性”命令，在打开的对话框中选中“Internet 协议版本 4（TCP/IPv4）”选项后单击“属性”按钮。

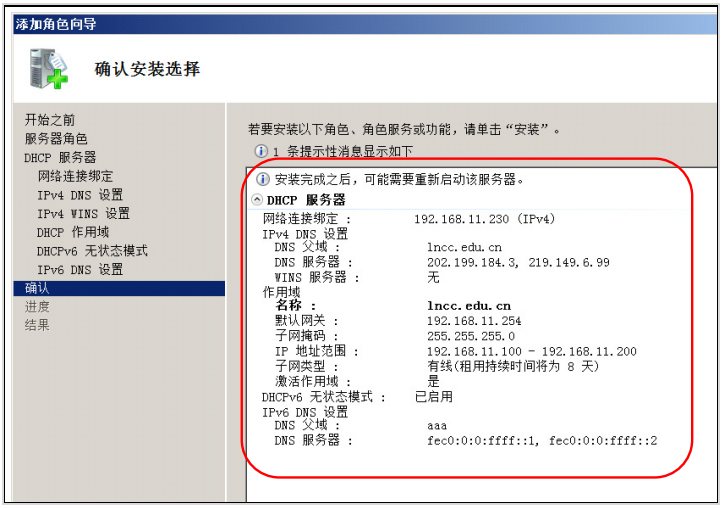


图 4.14　确认安装信息

图 4.15　安装成功

（3）在打开的“Internet 协议版本 4（TCP/IPv4）属性”对话框中，选择“自动获得 IP 地址”单选按钮，选择“自动获得 DNS 服务器地址”单选按钮，如图 4.16 所示。最后单击“确定”按钮。

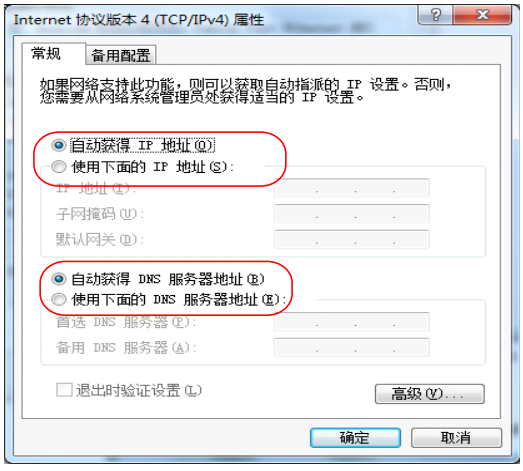


图 4.16　客户端 TCP/IP 参数

至此，服务器及客户端相关设置全部完成。在 DHCP 服务器正常运行的情况下，启动客户端主机将自动获得 IP 地址等网络参数。在客户端命令提示符方式下，使用 ipconfig/all 命令可以看到获得的 IP 地址等网络参数，如图 4.17 所示。

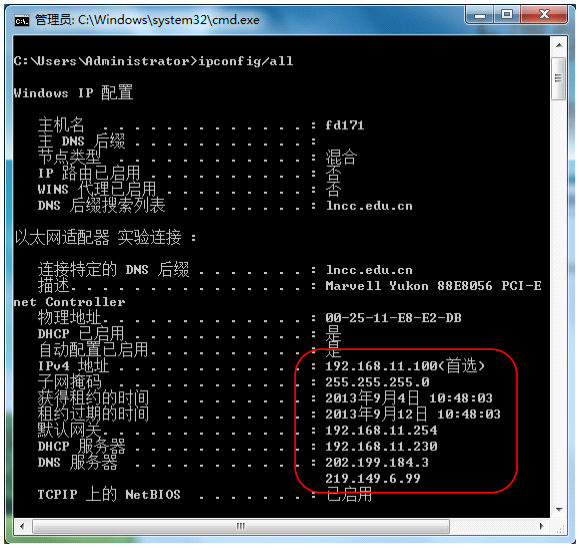


图 4.17　客户端自动获得的网络参数

4.4　演示实例

1. 背景描述

某公司办公局域网通过三层交换机将各部门及服务器连接，为了便于管理，各部门及服务器区分别位于不同的 VLAN 中。VLAN10 的 IP 地址为 192.168.10.1/24，VLAN20 的 IP 地址为 192.168.20.1/24，VLAN80 的 IP 地址为 192.168.80.1/24。在三层交换机 RUIJIE 上配置 DHCP 服务，为 VLAN10、VLAN20 客户端自动配置 IP 及其他网络参数，其中，服务器区中的本地 DNS 服务器的 IP 地址为 192.168.80.10，公网提供的 DNS 服务器 IP 地址为 202.199.184.1，域名为 lncc.edu.cn，每个 VLAN 中主机地址范围（200～254）不允许分配给客户端使用，如图 4.18 所示。

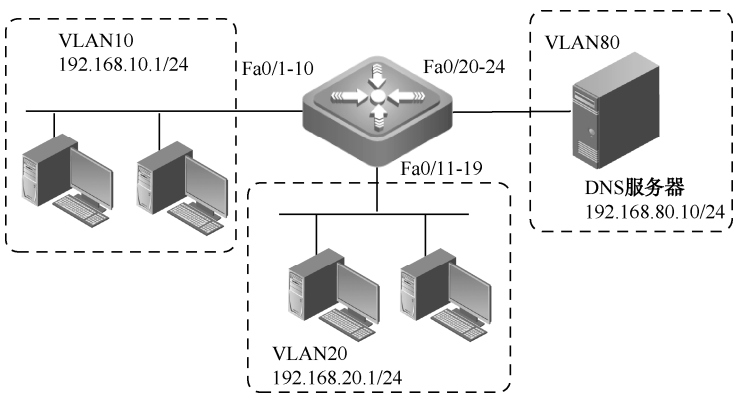


图 4.18　DHCP 操作演示

为了实现对 VLAN10 及 VLAN20 的客户端 IP 地址进行自动配置，需要在三层交换机上配置多地址池的 DHCP 服务器。服务器区 IP 地址采用静态配置。

2．操作步骤

第 1 步：创建 VLAN 并配置 IP 地址。

```
ruijie(config)#VLAN 10
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 20
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 80
ruijie(config-VLAN)#exit
ruijie(config)#interface range fastethernet 0/1-10
ruijie(config-if-range)#switchport mode access
ruijie(config-if-range)#switchport access VLAN 10
ruijie(config-if-range)#exit
ruijie(config)#interface range fastethernet 0/11-19
ruijie(config-if-range)#switchport mode access
ruijie(config-if-range)#switchport access VLAN 20
ruijie(config-if-range)#exit
ruijie(config)#interface range fastethernet 0/20-24
ruijie(config-if-range)#switchport mode access
ruijie(config-if-range)#switchport access VLAN 80
ruijie(config-if-range)#exit
ruijie(config)#interface VLAN 10
ruijie(config-if)#ip address 192.168.10.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#interface VLAN20
ruijie(config-if)#ip address 192.168.20.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#interface VLAN 80
ruijie(config-if)#ip address 192.168.80.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#
```

第 2 步：配置 DHCP 地址池及排除地址段。

```
ruijie(config)#service dhcp
ruijie(config)#ip dhcp pool VLAN10
ruijie(dhcp-config)#domain-name lncc.edu.cn
ruijie(dhcp-config)#network 192.168.10.0 255.255.255.0
ruijie(dhcp-config)#dns-server 192.168.80.10 202.199.184.1
ruijie(dhcp-config)#default-router 192.168.10.1
ruijie(dhcp-config)#exit
ruijie(config)#ip dhcp pool VLAN20
ruijie(dhcp-config)#domain-name lncc.edu.cn
```

```
ruijie(dhcp-config)#network 192.168.20.0 255.255.255.0
ruijie(dhcp-config)#dns-server 192.168.80.10 202.199.184.1
ruijie(dhcp-config)#default-router 192.168.20.1
ruijie(dhcp-config)#exit
ruijie(config)#ip dhcp excluded-address 192.168.10.200 192.168.10.254
ruijie(config)#ip dhcp excluded-address 192.168.20.200 192.168.20.254
```

第 3 步：测试。在客户端主机的本地连接的地址配置选项设置为“自动获取 IP 地址”，在主机命令行下使用 ipconfig/all 命令查看地址信息。

4.5 训练任务

1. 背景描述

某公司网络通过一个三层交换机连接技术部及市场部，三层交换机与路由器连接访问 Internet。要求在路由器上配置 DHCP 服务器功能为技术部及市场部提供 IP 地址、默认网关、DNS 服务器地址等参数，在三层交换机上配置 DHCP 服务器代理功能，如图 4.19 所示。

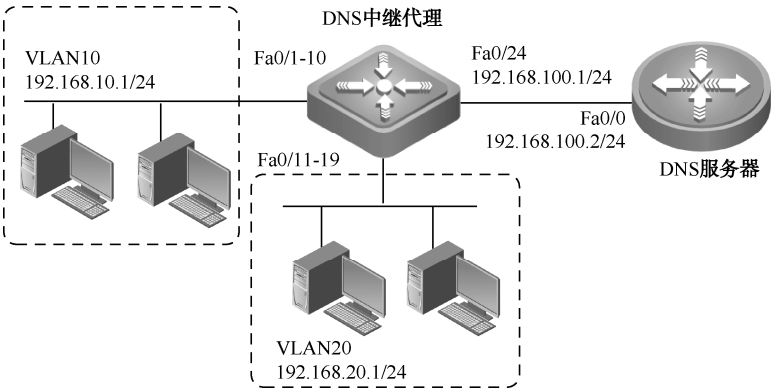


图 4.19 DHCP 技能训练

2. 操作提示

（1）交换机配置 DHCP 中继代理如下。

```
ruijie(config)#server dhcp
ruijie(config)#VLAN 10
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 20
ruijie(config-VLAN)#exit
ruijie(config)#VLAN 80
ruijie(config-VLAN)#exit
ruijie(config)#interface range fastethernet 0/1-10
ruijie(config-if-range)#switchport mode access
ruijie(config-if-range)#switchport access VLAN 10
ruijie(config-if-range)#exit
ruijie(config)#interface range fastether 0/11-19
```

```
ruijie(config-if-range)#switchport mode access
ruijie(config-if-range)#switchport access VLAN 20
ruijie(config-if-range)#exit
ruijie(config)#interface VLAN 10
ruijie(config-if)#ip address 192.168.10.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#interface VLAN20
ruijie(config-if)#ip address 192.168.20.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#no switchport
ruijie(config-if)#ip address 192.168.100.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#ip route 0.0.0.0 0.0.0.0 192.168.100.2
ruijie(config)#ip helper-address 192.168.100.2
ruijie(config)#
```

（2）路由器配置 DHCP 服务器中的地址池及排除地址段如下。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip address 192.168.100.2 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#service dhcp
ruijie(config)#ip dhcp pool VLAN10
ruijie(dhcp-config)#network 192.168.10.0 255.255.255.0
ruijie(dhcp-config)#dns-server 192.168.80.10 202.199.184.1
ruijie(dhcp-config)#default-router 192.168.10.1
ruijie(dhcp-config)#exit
ruijie(config)#ip dhcp pool VLAN20
ruijie(dhcp-config)#network 192.168.20.0 255.255.255.0
ruijie(dhcp-config)#dns-server 192.168.80.10 202.199.184.1
ruijie(dhcp-config)#default-router 192.168.20.1
ruijie(dhcp-config)#exit
ruijie(config)#ip dhcp excluded-address 192.168.10.1 192.168.10.10
ruijie(config)#ip dhcp excluded-address 192.168.20.1 192.168.20.10
ruijie(config)#ip route 192.168.10.0 255.255.255.0 192.168.100.1
ruijie(config)#ip route 192.168.20.0 255.255.255.0 192.168.100.1
ruijie(config)#
```

上述配置代码中 dns-server 地址可以根据实验环境自行决定。

练习题

1．选择题

（1）DHCP 协议使用的传输协议、源接口号、目标接口号是（　　）。

A．TCP　20　21　　B．TCP　67　68

C．UDP　20 21　　D．UDP　67　68

（2）DHCP 工作时，客户端与服务器间的通信报文有以下 4 种，正确的工作顺序是（　　）。

① 以广播方式发送 DHCPDISCOVER 报文

② 以单播方式发送 DHCPOFFER 报文

③ 以广播方式发送 DHCPREQUEST 报文

④ 以单播方式发送 DHCPACK 报文

A．①→②→③→④

B．③→②→①→④

C．①→③→②→④

D．②→①→③→④

2．简答题

（1）描述 DHCP 的工作过程。

（2）说明配置 DHCP 的命令及步骤。

项目 5　AAA 机制与 RADIUS 应用

访问控制是用来控制哪些人可以访问网络服务器及用户在网络上可以访问哪些服务的。身份验证、授权和记账（AAA）是进行访问控制的一种主要的安全机制。AAA 可以对单个用户（线路）或单个服务器动态配置身份验证、授权及记账类型。通过创建方法列表来定义身份验证、记账、授权类型，然后将这些方法列表应用于特定的服务或接口。

知识点、技能点

1. 了解 AAA 基本原理;
2. 掌握 AAA 特性及配置技能。

5.1　问题提出

某单位的计算机网络需要对登录该网络的用户进行集中管理，首先需要对用户身份进行验证，然后对不同的用户赋予不同的权限，必要时对通过身份验证的用户进行记账管理。

5.2　相关知识

5.2.1　AAA 概述

1. AAA 工作原理

AAA 是验证、授权和记账（Authentication Authorization and Accounting）的简称，它提供了对验证、授权和记账功能进行配置的一致性框架。AAA 可以对单个用户（线路）或单个服务器动态配置身份认证、授权及记账类型。通过创建方法列表来定义身份验证、记账、授权类型，然后将这些方法列表应用于特定的服务或接口。

AAA 以模块方式提供以下服务。

（1）验证（Authentication）。验证就是验证用户是否可获得访问权，可选择使用 RADIUS 协议或 Local 进行身份验证。身份验证是在允许用户访问网络和网络服务之前对其身份进行识别的一种方法。配置 AAA 就是定义一个身份验证方法列表并将其应用于各个接口的过程。

（2）授权（Authorization）。授权就是授权用户可使用哪些服务。AAA 授权是通过定义一系列的属性对来实现的，这些属性对描述了用户被授权执行的操作。这些属性对可以存放在网络设备上，也可以远程存放在 RADIUS 安全服务器上。

（3）记账（Accounting）。记账就是记录用户使用网络资源的情况。当 AAA 记账功能被启用时，网络设备便开始以统计记录的方式向 RADIUS 安全服务器发送用户使用网络资源的情况。每个记账记录都是以属性对的方式组成，并存放在安全服务器上的，这些记录可以通过专门的软件进行读取分析，从而实现对用户使用网络资源的情况进行记账、统计、跟踪。

实际上，除了 AAA 访问控制方法之外，对网络进行接入控制的方法还有很多，如本地用户名身份验证、线路密码身份验证等。这些安全验证与 AAA 相比，区别在于它们提供对网络保护的程度不一样，AAA 提供更高级别的安全保护。

使用 AAA 有以下优点。

第一，灵活性和可控制性强。

第二，可扩充性。

第三，标准化验证。

第四，多个备用系统。

应用 AAA 的基本拓扑结构如图 5.1 所示。

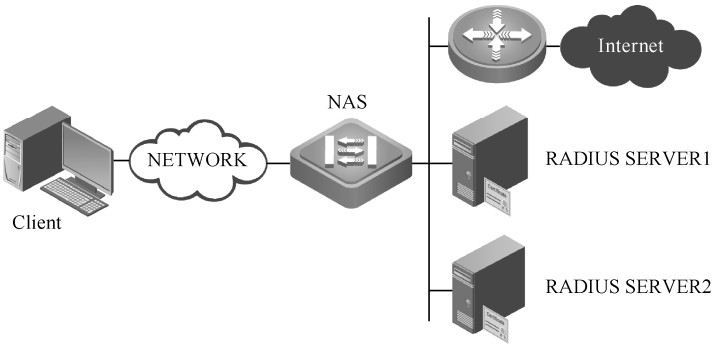


图 5.1　AAA 基本拓扑结构

当用户需要接入网络时，首先向 NAS（Netwoerk Access Server，网络接入服务器）发起请求。NAS 是一个提供网络接入或网络服务的设备（如交换机、路由器或服务器）。NAS 收到用户的接入请求后，将用户的请求发送至验证服务器（可以是一台或多台 RADIUS 服务器）。如图 5.1 所示为多台验证服务器时，NAS 设备会将用户请求信息首先发送给主验证服务器，由主验证服务器对用户身份进行验证，然后将验证结果（接受或拒绝）信息反馈给 NAS 设备，NAS 设备依据反馈信息对用户进行相应操作（接受或拒绝）。

当主验证服务器由于某种原因（如线路故障、设备故障等）导致响应超时时，NAS 设备将验证请求信息发送至备份验证服务器，由备份验证服务器进行用户身份验证。同时，如果 NAS 设备和验证服务器上配置了授权、记账功能，那么用户通过验证后将进行相应的授权和记账工作。

2．配置 AAA 过程

配置 AAA 过程，分为以下 3 步。

第 1 步：在 NAS 设备上启动 AAA 功能、定义验证方法列表、应用验证方法列表等。

第 2 步：配置 RADIUS 相关参数，包含指定 RADIUS 服务器的 IP 地址、NAS 设备与 RADIUS 服务器之间的密码等参数。必要时还需要搭建 RADIUS 服务器，配置服务器 IP 地址、创建用户、密码等。如果需要本地验证，还要建立本地用户数据库等。

第 3 步：配置授权、记账等相关参数。

5.2.2　配置 AAA 的验证功能

在 NAS 设备上，配置验证功能需要如下步骤。

第 1 步：启用 AAA 安全模型。

第 2 步：定义身份验证方法列表。

第 3 步：应用身份验证方法列表。

1. 启用 AAA 安全模型

为了使用 AAA 安全特性，必须首先在 NAS 设备上启用 AAA 安全模型。启用 AAA 安全模型的命令格式如下：

使用该命令的 no 选项可以关闭 AAA 安全服务，如 ruijie(config)#no aaa new-model。默认情况下，AAA 安全服务是关闭的。

例如：启用 AAA 安全服务。

```
ruijie(config)#aaa new-model
```

例如：关闭 AAA 安全服务。

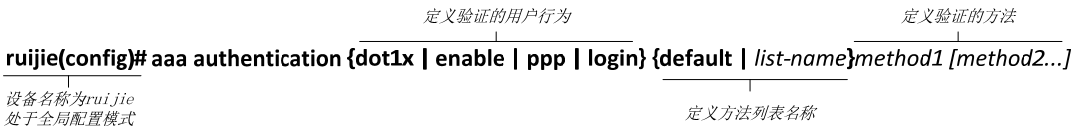
```
ruijie(config)#no aaa new-model
```

2. 定义身份验证方法列表

为了使用 AAA 身份验证功能，首先，需要定义一个身份验证方法列表，然后在特定的接口或服务上应用已定义的身份验证方法列表。除默认的身份验证方法列表外，对于已定义的身份验证方法列表，必须将其应用于特定的接口或服务时才生效。

身份验证方法列表定义了用户进行身份验证时的多种验证方法及验证顺序，这样确保在第一种方法失败时及时启用备份验证方法进行用户身份验证。NAS 设备首先使用方法列表中的第一种方法验证用户的身份，如果该方法无应答（服务器无响应，不包括用户信息非法情况），将选择方法列表中的下一种方法。 这个过程一直持续下去，直到成功地通过验证或用完方法列表中的验证方法。如果用完方法列表中的验证方法而还没有成功地通过验证时，则身份验证宣告失败。

定义用户身份验证方法列表的命令格式如下：



其中，定义验证的用户行为说明如下。

dot1x：定义针对 IEEE802.1x 接入用户进行身份验证。

enable：定义针对用户终端登录到 NAS 上的命令行界面（CLI）以后，需要提升 CLI 执行权限时进行用户身份验证，如从用户模式进入特权模式。

ppp：定义针对 PPP 拨号用户进行身份验证。

login：定义针对用户终端登录到 NAS 上的命令行界面，在登录时进行身份验证，如 telnet。

定义方法列表名称说明如下。

default：定义默认方法列表，默认情况下，默认方法列表应用到所有的接口、线路或服务。

list-name：定义用户身份验证的方法列表名称，可以通过该名称引用该方法列表。

定义验收方法及顺序说明如下。

Method1、Method2 等为下述关键字 ***group radius***、***guoup*** *group-name*、*local*、*none* 之一，实际使用过程中，可根据需要进行选用。

group radius：定义使用所有的 RADIUS 服务器进行验证。

group *group-name*：定义使用 RADIUS 服务器组中的服务器进行验证。

local：定义使用本地用户数据库进行验证。当配置 local 参数使用本地用户数据库进行身份验证时，需要使用 **username** *username* **password** *password* 命令预先在本地创建用户信息数据库。

none：定义不验证。此参数可以作为最后的备用验证方法，如果由于网络线路或设备故障导致无法正常进行验证时，在验证列表中的最后一步可以使用 none 不对用户进行验证，但当配置 none 后，在其后的配置其他验证方法将失效。

使用该命令的 no 选项删除用户身份验证的方法列表，如 ruijie(config)# no aaa authentication {dot1x | enable | ppp | login} {default |list-name}。

例如：将身份验证方法列表仅应用于线路 2。

```
ruijie#configure terminal
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password key123
ruijie(config)#aaa authentication login test group radius local
ruijie(config)#line vty 2
ruijie(config-line)#login    authentication test
```

在上述例子中，定义身份验证方法列表名称为 test，对 login 本地登录行为进行身份验证，验证方法有先使用（group radius）RADIUS 服务器验证 ，如果失败则使用（local）本地用户数据库验证。本地验证的用户名为 ruijie，密码为 key123。

例如：删除前面定义的名称为 test 的身份验证方法列表。

```
ruijie(config)#no aaa authentication login test
```

3．应用身份验证方法列表

定义的身份验证方法列表必须应用于接口或服务上才能生效。不同的身份验证方法列表的应用分别介绍如下。

（1）应用 login 验证方法列表。在很多情况下，用户需要通过 Telnet 方式来访问网络设备（NAS），一旦建立了这种连接，就可以远程配置 NAS 设备。为了防止未经授权的用户登录网络设备，需要对用户远程登录行为进行身份验证。

在预先定义了身份验证方法列表后，将身份验证方法列表应用到 login 登录行为上，实现通过 AAA 方式对远程登录的用户进行身份验证。

将身份验证方法列表应用到 login 登录行为上的命令格式如下：

定义的方法列表名称

ruijie(config-line)#login authentication {default | *list-name*}

设备名称为ruijie
处于线路配置模式

其中，*default* 为默认的验证方法列表名称；*list-name* 为定义的方法列表名称。

使用该命令的 no 选项删除应用的身份验证方法列表，如 ruijie(config-line)#no login authentication

{default | list-name}。

例如：如图 5.2 所示，在交换机线路 line0-4 上对 Telnet 登录的用户进行 AAA 本地身份验证。

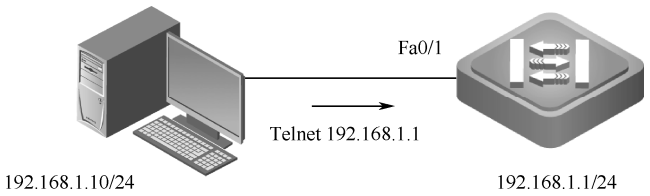


图 5.2　配置 Telnet 验证实例

交换机配置代码如下：

```
ruijie#configure terminal
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password key123
ruijie(config)#aaa authentication login test local
ruijie(config)#interface VLAN 1
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
ruijie(config-if)#exit
ruijie(config)#line vty 0 4
ruijie(config-line)#login authentication test
ruijie(config-line)#end
ruijie#
```

配置上述代码后，确认计算机的 IP 地址为 192.168.1.10 后，通过 ping 192.168.1.1 命令验证链路连通性。最后在计算机上执行如下 Telnet 命令。

```
C:\>telnet 192.168.1.1
username：ruijie
Password：key123
ruijie>
```

例如：删除应用在线路 line0-4 上的身份验证方法列表 test。

```
ruijie(config)#line vty 0 4
ruijie(config-line)#no login authentication test
```

（2）应用 PPP 验证方法列表。PPP 协议是提供在点到点链路上承载网络层数据包的一种链路层协议。在很多情况下，用户需要通过异步或 ISDN 拨号访问 NAS（网络访问服务器），一旦建立了这种连接，将启动 PPP 协商，为了防止未经授权的用户访问网络，PPP 在协商过程中要对拨号用户进行身份验证。

为了进行 PPP 用户身份验证，需要将前面定义的验证方法列表应用到 PPP 接口上。具体应用的命令格式如下：

验证方法

ruijie(config-if)# ppp authentication {chap | pap} ***{default | list-name}***

路由器名称为ruijie
处于接口配置模式

验证方法列表名称

其中，***default*** 为默认的验证方法列表名称；***list-name*** 为定义的方法列表名称。

使用该命令的 no 选项可以删除应用的方法列表，如 ruijie(config-if)#no ppp authentication chap {chap|pap} {default|list-name}。

例如：在接口 Serial 2/0 上对 PPP 协议用户进行 AAA 身份验证，首先进行 RADIUS 验证，如果失败则进行本地验证。

```
ruijie#configure terminal
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password key123
ruijie(config)#aaa authentication ppp test group radius local
ruijie(config)#interface serial 2/0
ruijie(config-if)#encapsulation ppp
ruijie(config-if)#ppp authentication pap test
ruijie(config-if)#exit
ruijie(config)#
```

例如：删除前面应用于接口 Serial 2/0 上名称为 test 的身份验证方法列表。

```
ruijie(config)#interface serial 2/0
ruijie(config-if)#encapsulation ppp
ruijie(config-if)#no ppp authentication pap test
```

（3）应用 Enable 验证方法列表。

如果登录到命令行界面后，由于初始权限过低，不能执行某些命令，则可以使用 enable 命令来提升权限，从用户模式进入到特权模式。为了防止网络未经授权的访问，在提升权限的时候，需要进行身份验证，即 Enable 验证。

前面已经介绍了定义 Enable 用户方法列表时采用如下命令：

定义验证的方法

ruijie(config)# aaa authentication enable default *method1 [method2...]*

设备名称为ruijie
处于全局配置模式

Enable 验证方法列表全局只能定义一个，因此不需要定义验证方法列表的名称；关键字 method 指的是实际验证的方法。Enable 验证方法列表定义以后立即自动生效。此后，在特权模式下执行 enable 命令的时候，如果要切换的级别比当前级别要高，则会提示进行验证。如果要切换的级别小于或等于当前级别，则直接切换，不需要进行验证。

如果采用本地验证的话，还需要建立本用户数据库，供验证使用。建立用户名称及密码的命令格式如下：

定义用户名称

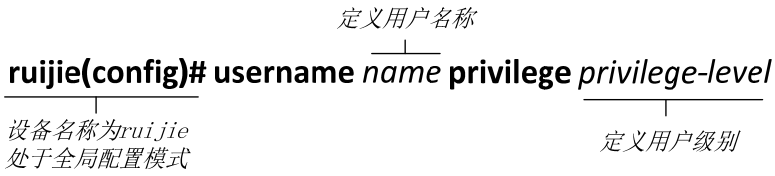
ruijie(config)# username *name* **password** *encrypted-password*

设备名称为ruijie
处于全局配置模式

定义用户密码

其中，*name* 为 用户名称；*encrypted-password* 为用户密码。

在默认情况下，系统只有两个受口令保护的授权级别：普通用户级别（1 级）和特权用户级别（15 级）。但是用户可以为每个模式的命令划分 15 个授权级别。通过定义用户级别，可以限制用户使用命令的权限，默认情况下用户处于 1 级。设置用户所处级别的命令格式如下：



其中，*name* 为 用户名称；*encrypted-password* 为 *privilege-level* 为用户级别。

例如：定义用户 user1 的密码为 pass1，处于 14 级。

```
ruijie(config)#username user1 password pass1
ruijie(config)#username user1 privilege 14
```

不同级别中包含不同的命令集合，要将一条命令的执行权限授予一个命令级别，可以在全局配置模式下使用 privilege 命令，其命令格式如下：

其中，*mode* 为要授权的命令所属的 CLI 命令模式；[all]为将指定命令的所有子命令的权限变为相同的权限级别；**level** *level* 为指定命令或子命令的运行权限级别，级别的范围为 0～15；**reset** 为将命令的执行权限恢复为默认级别；*command-string* 为要授权的命令字符串。

使用该命令的 no 选项可以将一条命令的执行权限恢复为默认值，如 ruijie(config)#privilege config all reset username，将 username 命令恢复为 15 级。

例如：定义 Enable 身份验证方法列表，使得本地用户 user14、密码为 pass14 的用户级别为 14 级可以使用 username 命令。

第 1 步：启动 AAA，配置 14 级用户 user14 及密码 pass14、15 级用户 user15 及密码 pass15。

```
ruijie(config)#aaa new-model
ruijie(config)#aaa authentication enable default local
ruijie(config)#username user14 password pass14
ruijie(config)#username user15 password pass15
ruijie(config)#username user14 privilege 14
ruijie(config)#username user15 privilege 15
ruijie(config)#end
ruijie#exit
ruijie>enable 14
username：user14
password：pass14
ruijie#show privilege
Current privilege level is 14
```

从上述显示结果知道，用户 user14 处于 14 级别上。在全局配置模式下使用? 命令查看命令。

```
ruijie(config)#?
Configure commands:
…
track                    Object tracking configuration commands
udp-helper               UDP broadcast packet relay
```

```
virtual-group          Virtual group bundling configuration
VLAN                   Configure VLAN
```

从上述显示结果知道，14 级中不包含 username 命令。

第 2 步：进入 15 级。

```
ruijie#enable 15
username:user15
password:pass15
ruijie#show privilege
Current privilege level is 15
ruijie#config
ruijie(config)#?
Configure commands:
…
track                  Object tracking configuration commands
udp-helper               UDP broadcast packet relay
username               Establish User Name Authentication
virtual-group            Virtual group bundling configuration
VLAN                   Configure VLAN（4）应用 802.1X 标准验证方法列表
```

从上述显示结果知道，15 级中不包含 username 命令。

第 3 步：设置 username 命令级别为 14 级。

```
ruijie(config)#privilege config all level 14 username
ruijie(config)#end
ruijie#disable 14          /将当前级别降为 14 级
ruijie#show privilege
Current privilege level is 14
```

从上显示结果看出，当前处于 14 级。

```
ruijie(config)#?
Configure commands:
…
track                  Object tracking configuration commands
udp-helper               UDP broadcast packet relay
username               Establish User Name Authentication
virtual-group            Virtual group bundling configuration
VLAN                   Configure VLAN
```

从上述显示结果知道，14 级中包含 username 命令。

（4）应用 dot1x 验证方法列表。IEEE 802.1X（Port-Based Network Access Control）是一个基于接口的网络访问控制标准，为 LAN 提供点对点式的安全接入，提供一种对连接到局域网设备的用户进行身份验证的手段。其详细描述见后续内容。在全局模式下使用如下命令应用身份验证方法列表：

定义的方法列表名称

ruijie(config-if)#dot1x authentication {default | *list-name*}

设备名称为ruijie
处于端口配置模式

其中，*default* 为默认的验证方法列表名称；*list-name* 为用户自定义的验证方法列表名称。

例如：在接口上应用身份验证方法列表并先进行 group radius 验证，若失败则进行本地验证。

```
ruijie#configure terminal
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password key123
ruijie(config)#aaa authentication dot1x list1 group radius local
ruijie(config)#interface fastEthernet0/1
ruijie(config-if)#dot1x authentication list1
ruijie(config-if)#end
ruijie#
```

5.2.3　RADIUS 服务

1．RADIUS 概述

RADIUS 是远程验证拨号用户服务（Remote Authentication Dial-In User Service）的简称，它是一种分布式的客户机/服务器系统，与 AAA 配合对试图连接的用户进行身份验证，防止未经授权的访问。在 RGOS 的实现中，RADIUS 客户端运行在设备或网络访问服务器（NAS）上，并向远程 RADIUS 服务器发出身份验证请求，远程 RADIUS 服务器包含了所有的用户身份验证和网络服务信息。由于 RADIUS 是一种完全开放的协议，很多系统如 UNIX、Windows 2000 等均将 RADIUS 服务器作为一个组件安装，因此 RADIUS 是目前应用最广泛的安全服务器。

RADIUS 协议的模型通常由三部分构成：RADIUS 服务器、Client（NAS 设备）和验证客户端，如图 5.3 所示。其中，NAS 设备作为 RADIUS 服务器的客户端，同时也作为验证客户端的服务器。当验证客户需要连接网络时，验证客户按照 NAS 要求格式将用户名称及密码发送给 NAS，NAS 再将验证信息发送给 RADIUS 服务器，RADIUS 服务器根据数据库中的记录信息返回验证结果。

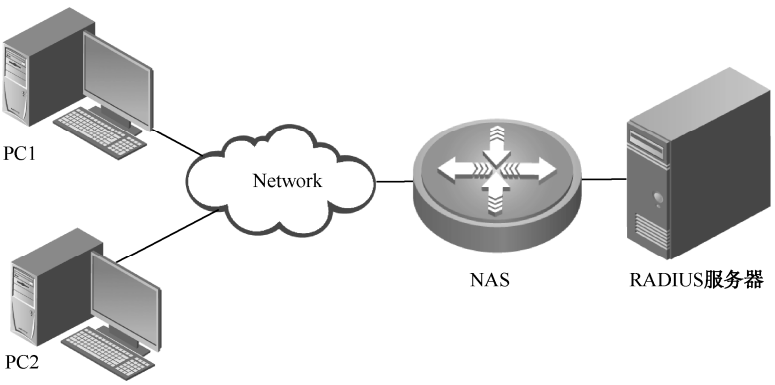


图 5.3　典型 RADIUS 网络结构

2．RADIUS 验证过程

如图 5.4 所示，RADIUS 验证过程如下。

（1）当用户接入 NAS 的时候，NAS 设备将为用户提供一个接入模式，使用户可以输入自己的账户和密码。这个过程通常以终端方式连接，即还没有进行数据链路层的协议连接，用户设备以远程终端方式挂接在 NAS 终端服务器上。

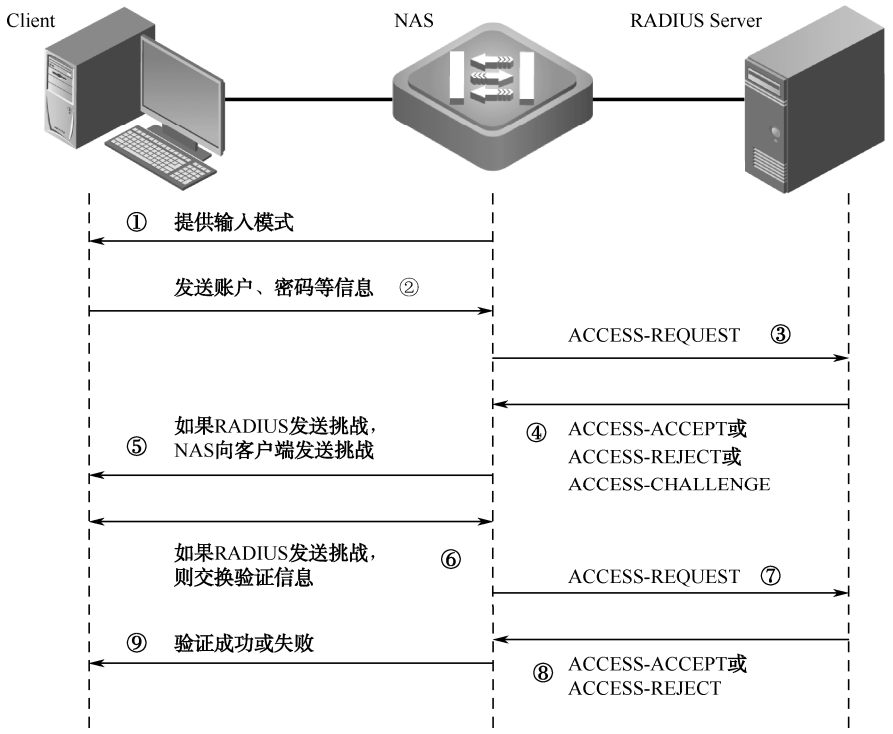


图 5.4 RADIUS 验证过程

（2）NAS 得到用户输入的账户和密码后，运行 RADIUS 客户程序生成一个 RADIUS 数据包，将相关的属性字段放置在其中，包括用户名、密码、NAS 的 IP 地址、NAS 的接口号等。在用户密码字段中，RADIUS 协议规定使用 MD5 算法进行加密。NAS 将这个数据包发送给 RADIUS 服务器，这个报文叫做“Access-Request”。

（3）Access-Request 报文通过网络从 NAS 传输至 RADIUS 服务器，并等待应答。

如果 NAS 在一段时间内未收到应答报文，则做出如下处理。

第 1 步：重试，重复将 Access-Request 报文发向同一个 RADIUS 服务器。

第 2 步：切换服务器，若超过重试次数，则切换 RADIUS 服务器，重复第 1 步。

第 3 步：连接出错，当前两步失败时，则认为 RADIUS 网络连接失败。

当 RADIUS 服务器接收到客户的请求报文（Access-Request），则做出如下处理。

第 1 步：首先进行客户验证。只有是合法客户的请求才做处理。RADIUS 服务器检查用户信息数据库，用户信息数据库包含合法客户的 IP 地址、对称密钥、客户类型等信息。

第 2 步：如果客户是合法的，RADIUS 服务器通过用户名的格式，判断用户是否为本地用户。若为本地用户，则 RADIUS 服务器根据用户名从用户数据库中查找记录。这条记录包含了允许用户访问的一些必要条件，如用户密码、被叫号码、主叫号码、NAS 的 IP 地址、NAS 接口号等。若是漫游客户，则开始一次代理会话，此时 RADIUS 服务器作为 NAS 设备的代理端将验证信息发送给相关验证服务器。

当用户满足条件时，RADIUS 服务器向客户（NAS）发送 Access-Accept 报文或挑战报文。

当用户不满足条件时，RADIUS 服务器向客户（NAS）发送 Access-Reject 报文。Access-Reject 报文包含字符串提示信息。NAS 设备接收到 Access-Reject 报文后，拒绝用户连接请求，并向用户表明拒绝访问的提示信息。

（4）如果发送的是 Access-Accept 报文，则由 NAS 设备向用户返回验证成功信息。
（5）如果发送的是挑战报文，则由 NAS 设备向验证客户端发起挑战。
（6）验证客户端接收到挑战报文后，验证客户端和 NAS 设备交换验证信息。
（7）NAS 设备向 RADIUS 服务器发送 Access-Request 报文，请求验证。
（8）RADIUS 服务器向 NAS 设备返回验证结果，发送 Access-Accept 或 Access-Reject 报文。
（9）NAS 设备向验证客户端返回验证成功或失败信息。

3．RADIUS 授权过程

RADIUS 授权和验证过程是协同进行的。当 RADIUS 服务器返回一个 Access-Accept 消息，该消息也包含了一系列的描述用户具有的会话属性（在 RADIUS 属性域内）。该信息为对 NAS 进行授权的信息。NAS 实现授权，并通知用户相应的成功或者失败信息，如图 5.5 所示。在接受消息中包含的授权数据一般为用户可以访问的服务，包括 Telnet、rlogin、PPP、SlIP、EXEC 等服务；连接参数，包括主机或者客户端 IP 地址、访问列表和用户超时值。

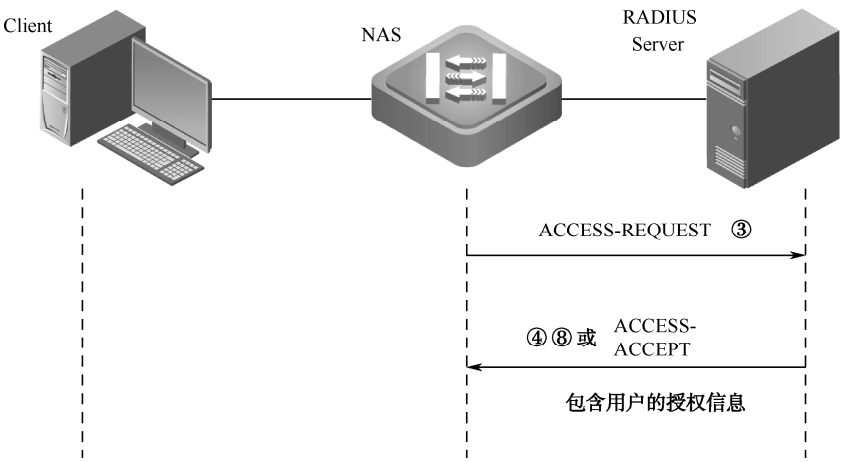


图 5.5　RADIUS 授权过程

4．RADIUS 计费过程

当一个客户端被配置为使用 RADIUS 计费时，在服务传送开始时，NAS 将产生一个计费开始请求包，用来描述被传送的服务类型及服务将传送到的用户，并向 RADIUS 服务器发送该计费开始请求包（Accounting-Request），RADIUS 服务器收到并成功记录该请求包后要回应一个计费响应（Accounting-Response）。

当用户断开连接时（连接也可以由接入服务器断开），NAS 向 RADIUS 服务器发送一个记费停止请求包，其中包含用户上网所使用网络资源的统计信息（上网时长、进/出的字节/包数等），RADIUS 服务器会返回对该计费停止请求包的确认，如图 5.6 所示。

5．RADIUS 基本配置

如果在部署 AAA 时采用 RADIUS 服务进行验证，则需要在 NAS（路由器或交换机）上进行 RADIUS 的相关配置，以告诉 NAS 当接收到客户的访问请求时将验证请求发送给哪个 RADIUS 服务器。

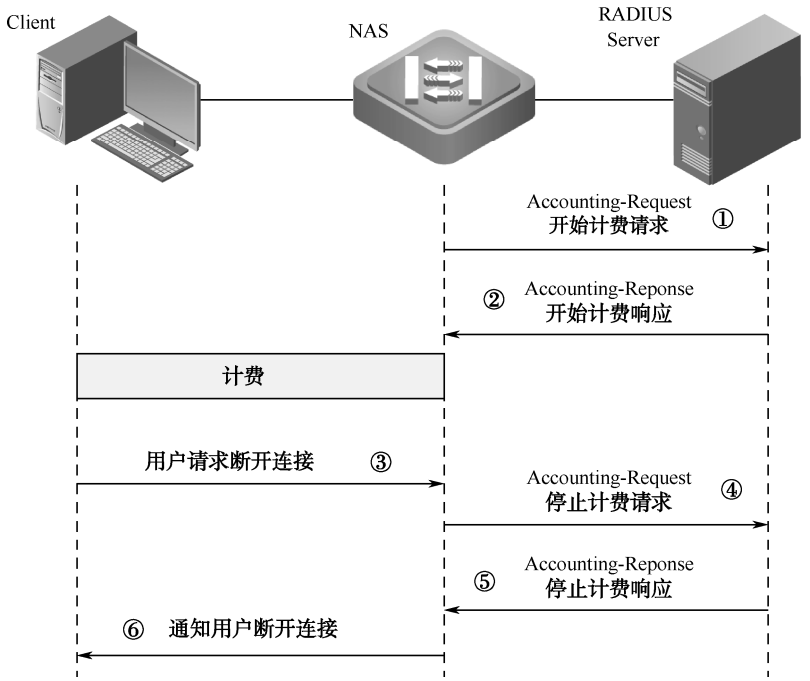


图 5.6　RADIUS 记账过程

（1）配置 RADIUS 服务器地址。要指定 RADIUS 服务器地址，在全局配置模式下执行如下命令：

配置RADIUS服务器地址命令　　RADIUS服务器IP地址　　指定记账UDP接口号

ruijie(config)# radius-server host *{hostname | ip-address}* **[auth-port** *port-number***] [acct-port** *port-number***]**

设备名称为ruijie处于全局配置模式　　RADIUS服务器DNS名称　　指定验证UDP接口号

其中，*hostname* 为RADIUS 服务器的 DNS 名称；*ip-address* 为RADIUS服务器的 IP 地址；*port-number* 为 RADIUS 身份验证的 UDP 接口号，如果设置为 0，则该主机不进行身份验证，默认情况下，RADIUS 服务器的验证和授权接口号为 UDP1812；*port-number* 为 RADIUS 记账的 UDP 接口号，如果设置为 0，则该主机不进行记账。默认情况下，RADIUS 服务器的记账接口号为 UDP1813。

提示：*使用该命令的 no 形式删除指定的 RADIUS 安全服务器主机，如 ruijie(config)# no radius-server host {ip-address}。默认情况下，没有指定的 RADIUS 服务器。*

可以使用此命令添加多个 RADIUS 服务器，当一个 RADIUS 服务器不可用时将使用下一个配置的 RADIUS 服务器，NAS 将按照配置的顺序进行查找。

例如：配置一个 RADIUS 安全服务器 IP 地址为 192.168.12.1。

```
ruijie(config)#radius-server host 192.168.12.1
```

例如：删除刚刚配置的 IP 地址为 192.168.12.1 的 RADIUS 服务器地址。

```
ruijie(config)#no radius-server host 192.168.12.1
```

（2）配置 RADIUS 服务器密码。为了保证 NAS 设备与 RADIUS 服务器之间通信的安全，可以配置 RADIUS 服务器验证密码，此密码用来提供 NAS 设备与 RADIUS 服务器之间通信的安全性。为了配置 RADIUS 服务器密码，需要在全局配置模式下执行如下命令：

其中，*text-string* 为服务器密码字符串。

提示：*使用该命令的 no 形式取消指定的验证密码，如 ruijie(config)# no radius-server key。默认情况下，没有指定任何验证密码。*

此命令设置的密码必须与 RADIUS 服务器中设置的密码完全匹配，否则 RADIUS 服务器将丢弃验证报文。

例如：定义 RADIUS 安全服务器的验证密码为 key123。

```
ruijie(config)#radius-server key key123
```

例如：取消刚刚定义的 RADIUS 服务器验证密码。

```
ruijie(config)#no radius-server key
```

（3）创建 RADIUS 服务器组。使用 radius-server host 命令配置完 RADIUS 服务器后，可以将多个 RADIUS 服务器放置到一个服务器组中。

在全局配置模式下，使用如下命令创建 RADIUS 服务器组：

其中，*name* 为服务器组的名称，目前不能为关键字“radius”和“tacacs+”。

提示：*使用该命令的 no 形式删除服务器组，如 ruijie(config)# no aaa group server {radius | tacacs+} name。*

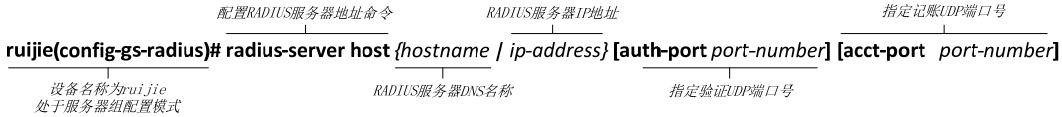
例如：创建一个配置名称为 ss 的 RADIUS 服务器组。

```
ruijie(config)#aaa group server radius ss
ruijie(config-gs-radius)#
```

例如：删除名称为 ss 的 RADIUS 服务器组。

```
ruijie(config)#no aaa group server radius ss
```

配置完此命令后，系统将进入服务器组视图中。在服务器组视图下使用如下命令将 RADIUS 服务器加入到服务器组中。



其中，*hostname* 为 RADIUS 安全服务器主机的 DNS 名称；*ip-address* 为 RADIUS 安全服务器主机的 IP 地址；*port-number* 为 RADIUS 身份验证的 UDP 接口号，如果设置为 0，则该主机不进行身份验证；*port-number* 为 RADIUS 记账的 UDP 接口号，如果设置为 0，则该主机不进行记账。

提示：*使用该命令的 no 形式删除指定的 RADIUS 安全服务器主机，如 ruijie(config-gs-radius)# no radius-server host {ip-address}。*

加入到服务器组中的 RADIUS 服务器必须是已经在全局配置模式下使用 radius-server host 命令定

义过的。

配置完RADIUS服务器后，最后还需要在AAA的验证方法列表中配置RADIUS的验证方法（**group radius** 或 **group** *group-name*），这样 NAS 设备才会将请求发送给配置的 RADIUS 服务器进行验证。当使用 RADIUS 服务器进行验证时，要确保 NAS 与 RADIUS 服务器之间数据可达，并且 RADIUS 服务器上已经启用 RADIUS 相关服务。

例如：将预先配置 IP 地址为 192.168.12.1、172.16.1.1 的 RADIUS 安全服务器加入到名称为 RS 的 RADIUS 服务器组中。

```
ruijie(config)#radius-server host 192.168.12.1
ruijie(config)#radius-server host 172.16.1.1
ruijie(config)#aaa group server radius RS
ruijie(config-gs-radius)#radius-server host 192.168.12.1
ruijie(config-gs-radius)#radius-server host 172.16.1.1
ruijie(config-gs-radius)#
```

6．RADIUS 高级配置

（1）配置 RADIUS 超时时间。当 NAS 向 RADIUS 服务器发出请求后，如果在超时时间（timeout）内没有收到 RADIUS 服务器的响应，则 NAS 会重传请求。在全局配置模式下可以通过如下命令设置 NAS 等待 RADIUS 服务器的响应超时时间：

配置RADIUS服务器超时时间命令

ruijie(config)# radius-server timeout *seconds*

设备名称为ruijie
处于全局配置模式

定义的超时时间

其中，*seconds* 为超时时间（单位为秒），可设置的值范围为 1～1000 秒，默认的超时时间为 5 秒。

提示：*使用该命令的 no 形式恢复默认值，如 ruijie(config)# no radius-server timeout。*

例如：将 RADIUS 服务器超时时间定义为 10 秒。

```
ruijie(config)#radius-server timeout 10
```

例如：将 RADIUS 服务器超时时间恢复为默认值 5 秒。

```
ruijie(config)#no radius-server timeout
```

（2）配置 RADIUS 重传次数。当 NAS 在超时时间内没有收到 RADIUS 服务器的响应后，将向 RADIUS 服务器重传请求报文。在全局模式下可以通过如下命令设置重传次数：

配置RADIUS服务器重传次数命令

ruijie(config)# radius-server retransmit *retries*

设备名称为ruijie
处于全局配置模式

定义的重传次数

其中，*retries* 为 RADIUS 尝试重传次数。

提示：*使用该命令的 no 形式恢复默认重传次数，如 ruijie(config)#no radius-server retransmit。默认的重传次数为 3 次。*

例如：指定 RADIUS 服务器重传次数为 6。

```
ruijie(config)#radius-server retransmit 6
```

例如：恢复 RADIUS 服务器重传次数为默认值 3。

```
ruijie(config)#no radius-server retransmit
```

（3）配置 RADIUS 服务器的死亡时间。当 NAS 向 RADIUS 服务器发送请求报文的重传次数已经达到设置最大值后仍没有收到 RADIUS 服务器的响应，则认为该 RADIUS 服务器已“死亡”。使用如下命令可以设置 RADIUS 服务器的“死亡”时间，在这个时间间隔内，NAS 将不会再向该 RADIUS 服务器发送请求报文，该服务器将被立即跳过。

配置RADIUS服务器死亡时间命令

ruijie(config)# radius-server deadtime *minutes*

设备名称为ruijie
处于全局配置模式　　*定义的死亡时间*

其中，*minutes* 为死亡时间（单位为分钟），可设置的值范围为 1～1000 秒。

提示：*使用该命令的 no 形式，恢复默认值，如 ruijie(config)# no radius-server deadtime。默认的死亡时间为 5 分钟。*

例如：将 RADIUS 服务器的死亡时间定义为 10 分钟。

```
ruijie(config)#radius-server deadtime 10
```

例如：将刚刚定义 RADIUS 服务器的死亡时间恢复为默认值 5 分钟。

```
ruijie(config)#no radius-server deadtime
```

7．应用案例

某公司网络中，需要对远程登录到 NAS 设备上的用户进行 AAA 验证，验证方式首先使用 RADIUS 进行验证，如果 RADIUS 无法访问，则进行本地验证，如图 5.7 所示。

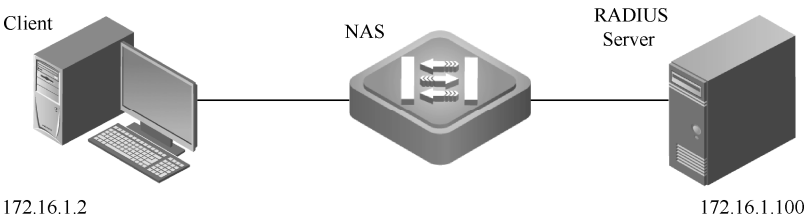


图 5.7　AAA 与 RADIUS 配置

NAS 设备配置代码如下：

```
ruijie#configure terminal
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password 123456
ruijie(config)#aaa authentication login test group radius local
ruijie(config)#radius-server host 172.16.1.100
ruijie(config)#radius-server key key123
ruijie(config)#line vty 0 4
ruijie(config-line)#login authentication test
ruijie(config-line)#end
ruijie(config)#
```

其中，123456 为用户 ruijie 的登录密码，key123 为 NAS 与 RADIUS 服务器之间的验证密码。

5.3 扩展知识

5.3.1 配置 AAA 的授权功能

AAA 授权功能使管理员能够对用户可使用的网络服务或权限进行控制。启用 AAA 授权服务以后，网络设备通过本地或远程服务器中的用户配置文件信息对用户的会话进行配置。完成授权以后，该用户只能使用配置文件中允许的网络服务或只具备许可的权限。

目前锐捷设备支持以下 AAA 授权类型。

（1）Exec 授权：针对的是用户终端登录到 NAS 上的 CLI 界面时，授予用户终端的权限级别（分为 0～15 级）。

（2）Command（命令）授权：针对的是用户终端登录到 NAS 上的 CLI 界面以后，针对具体命令的执行授权。

（3）Network（网络）授权：针对的是授予网络连接上的用户会话可使用的网络服务。

在配置 AAA 授权以前，必须完成下述任务。

第 1 步：启用 AAA 服务，前面已经有过详细介绍。

第 2 步：（可选）配置 AAA 验证。授权一般是在用户通过验证之后进行，但在某些情况下，没有经过验证也可以单独授权。

第 3 步：（可选）配置安全协议参数。如果需要使用安全协议进行授权，则需要配置安全协议参数。

第 4 步：如果需要使用本地授权，则需要使用 username 命令定义用户权限。

5.3.2 配置 AAA 的记账功能

AAA 记账功能够跟踪用户使用的服务和网络资源。启用记账功能以后，网络访问服务器或设备实时地以属性对的方式将用户访问网络的情况发送给 RADIUS 服务器，这些信息包括报文的个数及字节数、IP 地址、用户名等。可以使用一些分析软件对这些数据进行分析，实现对用户的活动进行计费、审计及跟踪等功能。

锐捷设备目前支持以下记账类型。

（1）Exec 记账。

（2）Command（命令）记账。

（3）Network（网络）记账。

其中，Exec 记账针对的是用户终端登录到 NAS 上的命令行界面（CLI），在登入和退出时分别进行记账；命令记账针对的是用户终端登录到 NAS 上的 CLI 界面以后，记录其具体执行的命令；网络记账针对的与网络连接上的用户会话有关的信息。

在配置 AAA 记账功能之前，必须完成以下任务。

第一，启用 AAA 安全服务。

第二，定义安全协议参数。进行记账，需要配置安全协议参数。

第三，（可选） 配置 AAA 验证。 某些记账需要在用户通过验证之后才进行（如 Exec 记账），其他情况下，没有经过验证也可以进行记账。

5.4　演示实例

1．任务描述

某公司网络拓扑如图 5.8 所示。为了便于网络管理工作，在网络设备上启用了远程登录功能。但是出于安全性考虑，管理员计划对远程登录用户进行 AAA 验证，提供网络安全性。

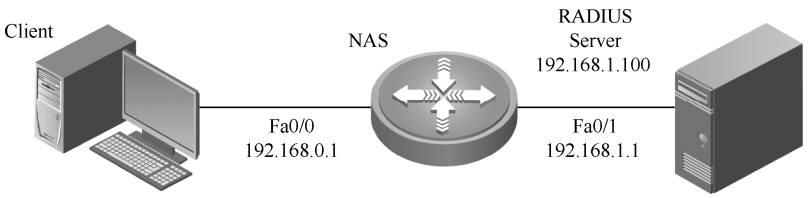


图 5.8　某公司网络拓扑图

2．操作步骤

为了实现远程登录用户的安全访问，可以利用网络设备的 AAA 验证功能。同时出于可用性考虑，在 AAA 验证方法列表中配置两种验证方法，RADIUS 服务器验证和本地验证，优先选择 RADIUS 服务器验证，当 RADIUS 服务器没有响应时选择本地验证方法。

具体配置如下。

第 1 步：配置路由器接口 IP 地址。

```
ruijie(config)#interface Fastethernet 0/0
ruijie(config-if)#ip address 192.168.0.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#exit
ruijie(config)#interface Fastethernet 0/1
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#exit
ruijie(config)#
```

第 2 步：配置路由器 AAA 验证方法列表。

```
ruijie(config)#aaa new-model
ruijie(config)#username ruijie password 123456
ruijie(config)#aaa authentication login test group radius local
ruijie(config)#radius-server host 192.168.1.100
ruijie(config)#radius-server key key123
```

第 3 步：在线路上应用 AAA 验证方法列表。

```
ruijie(config)#line vty 0 4
ruijie(config-line)#login authentication test
ruijie(config-line)#end
ruijie#
```

第 4 步：搭建 RADIUS 服务器。

可以启用相关的 RADIUS 应用软件，请自行配置 RADIUS 服务器的 IP 地址、与 NAS 连接的密码

key123，同时还要添加用户名称、用户登录密码等参数。

第 5 步：测试。

配置完成后，在 PC 上使用命令 telnet 192.168.0.1 远程登录到路由器上，同时在路由器上用 debug aaa 命令观察 AAA 验证调试信息。

5.5 训练任务

1. 背景描述

某公司总部与分公司之间通过 PPP 链路连接，为了提高接入安全性，公司要求各分公司通过 PPP 链路接入公司总部时都要进行身份验证，为了不影响路由器的性能，考虑通过 RADIUS 服务器进行 AAA 验证，如图 5.9 所示。

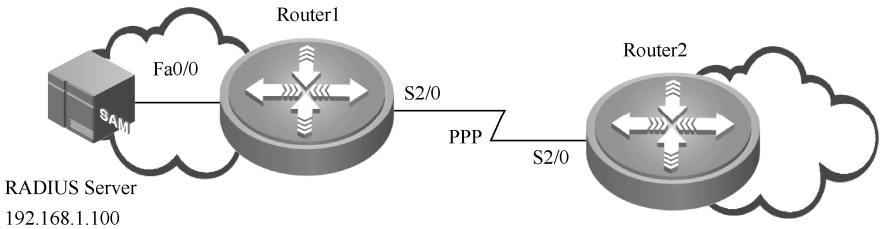


图 5.9　AAA 验证功能技能训练

2. 操作提示

在总公司接入服务器上配置 PPP 协议的 AAA 验证，当分公司请求连接时，发送用户名和密码给总公司路由器，总公司路由器发送给 AAA 验证服务器，验证通过则建立连接，否则连接失败。

主要配置代码如下。

第 1 步：在总公司路由器 Router1 上启动 AAA 验证功能、定义验证方法列表、设置 RADIUS 服务器 IP 地址及验证密码。

```
Router1(config)#aaa new-model
Router1(config)#aaa authentication ppp test group radius
Router1(config)#radius-server host 192.168.1.100
Router1(config)#radius-server key key123
```

第 2 步：在总公司路由器的接口上应用 AAA 验证。

```
Router1(config)#interface serial 2/0
Router1(config-if)#encapsulation ppp
Router1(config-if)#ppp authentication pap test
```

第 3 步：在客户端路由器上封装 PPP、发送验证用户名及密码。

```
Router2(config)#interface serial 2/0
Router2(config-if)#encapsulation ppp
Router2(config-if)#ppp pap sent-username ruijie password 0 key123
Router2(config-if)#
```

练习题

1．选择题

（1）下面说法是错误的是（　　）。

A．AAA 是 Authentication、Authorization and Accounting 的缩写词

B．AAA 是提供了对验证、授权和记账功能进行配置的一致性框架

C．AAA 必须使用 RADIUS 服务器的一种验证、授权、计费的安全模式

D．AAA 通过创建方法列表来定义身份验证、记账、授权类型，然后将这些方法列表应用于特定的服务或接口

（2）下面不属于 AAA 安全认证系统中的设备是（　　）。

A．RADIUS 服务器　　B．Client（NAS 设备）　C．验证客户端　　D．调制解调器

（3）若定义 telnet 登录行为进行验证，需要选择（　　）验证操作。

A．dot1x　　B．enable　　C．ppp　　D．login

（4）配置 RADIUS 服务器密钥用于（　　）加密。

A．NAS 与 RADIUS 服务器之间　　B．客户端与 NAS 之间

C．客户端与 RADIUS 服务器之间　　D．客户端与客户端之间

（5）默认情况下，RADIUS 服务器的认证和授权接口是（　　）。

A．TCP 1812　　B．UDP 1812　　C．TCP 1813　　D．UDP1813

（6）默认情况下，RADIUS 服务器的计费接口是（　　）。

A．TCP 1812　　B．UDP 1812　　C．TCP 1813　　D．UDP1813

2．简答题

（1）描述 AAA 能提供的主要服务有哪些。

（2）说明配置 AAA 的步骤。

（3）描述 RADIUS 协议模型通常由哪些部分构成。

（4）说明 RADIUS 验证过程。

项目 6　802.1X 安全访句控制

网络安全接入控制是网络安全的重要组成部分，在前面学习中我们采用安全接口、地址绑定、限制最大连接数等措施实现网络接入安全控制。但这些方法使用起来不够灵活，不能针对客户实施网络接入管理，本章介绍的 802.1X 技术可以解决此问题。

知识点、技能点

1. 了解 802.1X 验证体系构成及工作机制;
2. 掌握 802.1X 配置技能。

6.1　问题提出

某公司为了防止有公司外部的用户私自将计算机接入到公司网络中，造成公司机密信息的外泄，要求员工登录公司网络时进行身份验证，只有具有合法身份的用户才可以登录到公司网络中。

6.2　相关知识

1. 802.1X 概述

在局域网中，用户只要能接到网络设备上，不需要经过任何身份验证和授权即可直接使用。这样，一个未经授权的用户可以没有任何阻碍地通过连接到局域网的设备进入网络。随着局域网技术的广泛应用，特别是在运营网络的出现，对网络安全验证的需求已经提到了议事日程上。如何在简单、廉价的以太网技术的基础上，提供用户对网络或设备访问的合法性验证，已经成为业界关注的焦点。IEEE 802.1X 协议正是在这样的背景下提出的。

IEEE 802.1X（Port-Based Network Access Control）是一个基于接口的网络访问控制标准，为 LAN 提供点对点式的安全接入。这是 IEEE 标准委员会针对以太网的安全缺陷而专门制定的标准，能够在利用 IEEE 802 LAN 优势的基础上，提供一种对连接到局域网设备的用户进行身份验证的手段。

2. 802.1X 验证体系构成

IEEE 802.1X 标准定义了一种基于“客户-服务器”（Client-Server）模式实现了限制非法用户对网络的访问。客户端要访问网络必须先通过服务器的身份验证。在客户端通过身份验证之前，只有 EAPoL 报文（Extensible Authentication Protocol over LAN）可以在网络上通行。在身份验证成功之后，正常的数据流才可在网络上通行。应用 802.1X 的设备提供了 AAA 安全功能。

IEEE 802.1X 标准验证体系由恳请者、验证者、验证服务器三个角色构成，在实际应用中，三者分别对应为：工作站（Client）、设备（Network Access Server，NAS）、Radius-Server，如图 6.1 所示。

（1）恳请者。恳请者是最终用户所扮演的角色，一般是个人计算机。它请求对网络服务的访问，并对验证者的请求报文进行应答。恳请者必须运行符合 IEEE 802.1X 客户端标准的软件，目前最典型的就是 Windows XP 操作系统自带的 IEEE 802.1X 客户端支持，另外，锐捷也已推出符合该客户端

标准的 STAR Supplicant 软件。

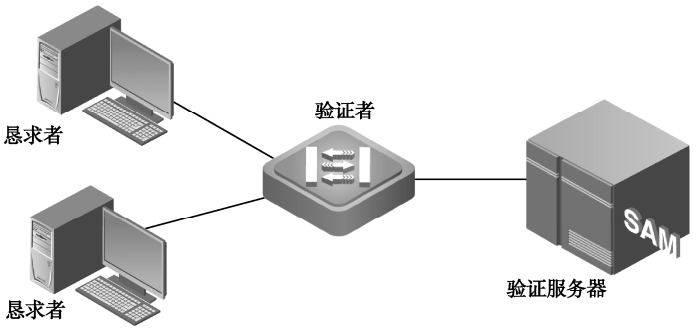


图 6.1　IEEE 802.1X 验证体系

（2）验证者。验证者一般为交换机等接入设备。该设备的职责是根据客户端当前的验证状态控制其与网络的连接状态。在客户端与服务器之间，该设备扮演着中介者的角色：从客户端要求用户名，核实从服务器端的验证信息，并且转发给客户端。因此，设备除了扮演 IEEE 802.1X 的验证者的角色外，还扮演 RADIUS Client 角色，因此我们把设备称作 Network Access Server（NAS），它要负责把从客户端收到的回应封装到 RADIUS 格式的报文并转发给 RADIUS Server，同时它要把从 RADIUS Server 收到的信息解释出来并转发给客户端。

扮演验证者角色的设备有两种类型的接口：受控接口（Controlled Port）和非受控接口（Uncontrolled Port）。连接在受控接口的用户只有通过验证才能访问网络资源；而连接在非受控接口的用户无须经过验证便可以直接访问网络资源。我们把用户连接在受控接口上，便可以实现对用户的控制；非受控接口主要是用来连接验证服务器，以便保证服务器与设备的正常通信，如图 6.2 所示。

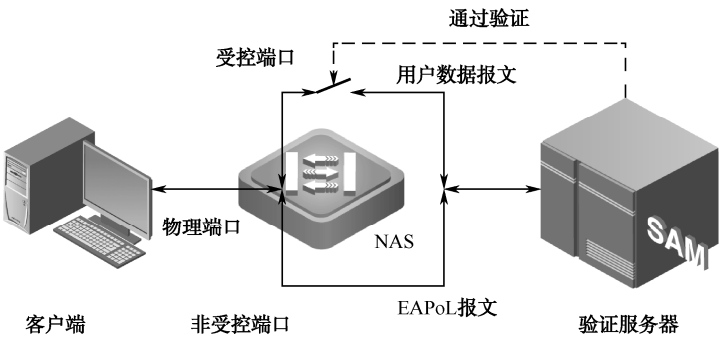


图 6.2　802.1X 验证接口

（3）验证服务器。验证服务器通常为 RADIUS 服务器，验证过程中与验证者配合，为用户提供身份验证服务。验证服务器保存了用户名及密码，以及相应的授权信息，一台服务器可以对多台验证者提供验证服务，这样就可以实现对用户的集中管理。验证服务器还负责管理从验证者发来的记账数据。锐捷网络科技公司实现 802.1X 的设备完全兼容标准的 Radius Server，如 Microsoft Windows 2000 Server 自带的 Radius Server 及 Linux 下的 Free Radius Server。

3. 802.1X 工作过程

恳请者和验证者之间通过 EAPoL 协议交换信息，而验证者和验证服务器通过 RADIUS 协议交换信息，通过这种转换完成验证过程，如图 6.3 所示。

图 6.3　802.1X 工作机制

在客户端与交换机之间，EAP（Extensible Authentication Protocol）报文直接被封装到 LAN 协议（如 Ethernet）中，形成 EAPoL 报文。在交换机与 RADIUS 服务器之间，EAP 协议报文被封装到 RADIUS 报文中，形成 EAPoRADIUS 报文。此外，在交换机与 RADIUS 服务器之间还可以使用 RADIUS 协议交互 PAP 和 CHAP 报文。交换机在整个验证过程中不参与验证，所有的验证工作都由 RADIUS 服务器完成。RADIUS 服务器可以使用不同的验证方式对客户端进行验证，如 EAP-MD5、PAP、CHAP 等。当 RADIUS 服务器对客户端身份进行验证后，将验证结果（接受或拒绝）返回给交换机，交换机根据验证结果决定受控接口的状态。

从验证方式来讲，802.1X 支持两种验证模式：EAP 中继模式和 EAP 终结模式，分别描述如下。

（1）EAP 中继模式。EAP 中继模式是 IEEE 802.1X 标准中定义的验证模式，交换机将 EAP 协议报文封装到 RADIUS 协议报文中，通过网络发送到 RADIUS 服务器。对于这种验证模式，需要 RADIUS 服务器支持 EAP 属性。EAP 中继模式的验证方式有 EAP-MD5、EAP-TLS（Transport Layer Security，传输层安全）等。

EAP 中继模式（EAP-MD5）验证过程如图 6.4 所示，具体描述如下。

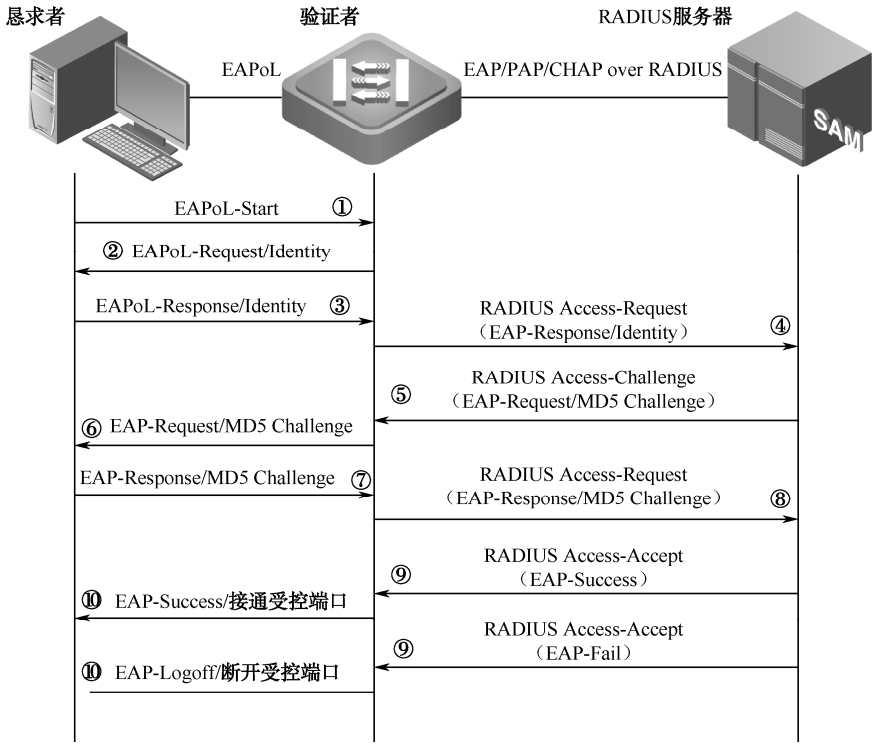


图 6.4　EAP 中继模式验证过程

① 客户端启动 802.1X 客户端软件，向交换机发送一个 EAPoL 报文，表示开始进行 802.1X 接入验证。

② 如果交换机接口启用了 802.1X 验证，将向客户端发送 EAP-Request/Identity 报文，要求客户端发送其使用的用户名（ID 信息）。

③ 客户端响应交换机发送的请求，向交换机发送 EAP-Response/Identity 报文，报文中包含客户端使用的用户名。

④ 交换机将 EAP-Response/Identity 报文封装到 RADIUS 的 Access-Request 报文中，通过网络发送给 RADIUS 服务器。

⑤ RADIUS 服务器收到交换机发送的 RADIUS 报文后，使用报文中的用户名信息在本地用户数据库中查找到对应的密码后，用随机生成的挑战值（MD5 Challenge）与密码进行 MD5 运算，产生一个 128bit 的散列值。同时 RADIUS 服务器也将此挑战值通过 RADIUS 的 Access-Challenge 报文发送给交换机。

⑥ 交换机从 RADIUS 报文中提取出 EAP 信息（其中包含挑战值），封装到 EAP-Request/MD5 Challenge 报文中发送给客户端。

⑦ 客户端使用报文中的挑战值与本地的密码也进行 MD5 运算，产生一个 128bit 的散列值，封装到 EAP-Response/MD5 Challenge 报文中发送给交换机。

⑧ 交换机将 EAP-Response/MD5 Challenge 报文中的信息封装到 RADIUS 的 Access-Request 报文中发送给 RADIUS 服务器。

⑨ RADIUS 服务器通过将收到的客户端的散列值与自己计算的散列值进行比较，如果相同则表示用户合法，验证通过，并返回 RADIUS Accept 报文，其中包含 EAP-Success 信息。

⑩ 交换机收到验证通过的信息后，将连接客户端的接口“开放”，并发送 EAP-Success 报文给客户端，以通过客户端验证通过。或者没有验证失败，交换机受控接口仍处于接口断开状态。

客户端可以通过发送 EAP-Logoff 报文通过交换机主动下线，终止验证状态。交换机收到 EAP-Logoff 报文后将接口“关闭”。从 EAP 中继模式的验证过程可以看出，交换机在整个验证过程中扮演着一个中间人的角色，对 EAP 报文进行透传。

（2）EAP 终结模式。EAP 终结模式即交换机将 EAP 信息终结，交换机与 RADIUS 服务器之间无须交互 EAP 信息，也就是说，RADIUS 服务器无须支持 EAP 属性。如果网络中的 RADIUS 服务器不支持 EAP 属性，则可以使用这种验证模式。

在 EAP 终结模式中可以使用 PAP 与 CHAP 验证方式，并且推荐使用 CHAP 验证方式，因为 PAP 使用明文传送用户名和密码信息。

EAP 终结模式 CHAP 方式验证过程，如图 6.5 所示，具体描述如下。

① 客户端启动 802.1X 客户端软件，向交换机发送一个 EAPoL 报文，表示开始进行 802.1X 接入验证。

② 如果交换机接口启用了 802.1X 验证，将向客户端发送 EAP-Request/Identity 报文，要求客户端发送其使用的用户名（ID 信息）。

③ 客户端响应交换机发送的请求，向交换机发送 EAP-Response/Identity 报文，报文中包含客户端使用的用户名。

④ 交换机将随机生成的挑战值（MD5 Challenge）封装到 EAP-Request/MD5 Challenge 报文中发送给客户端。

⑤ 客户端使用报文中的挑战值与本地的密码进行 MD5 运算，产生一个 128bit 的散列值，封装到 EAP-Response/MD5 Challenge 报文中发送给交换机。

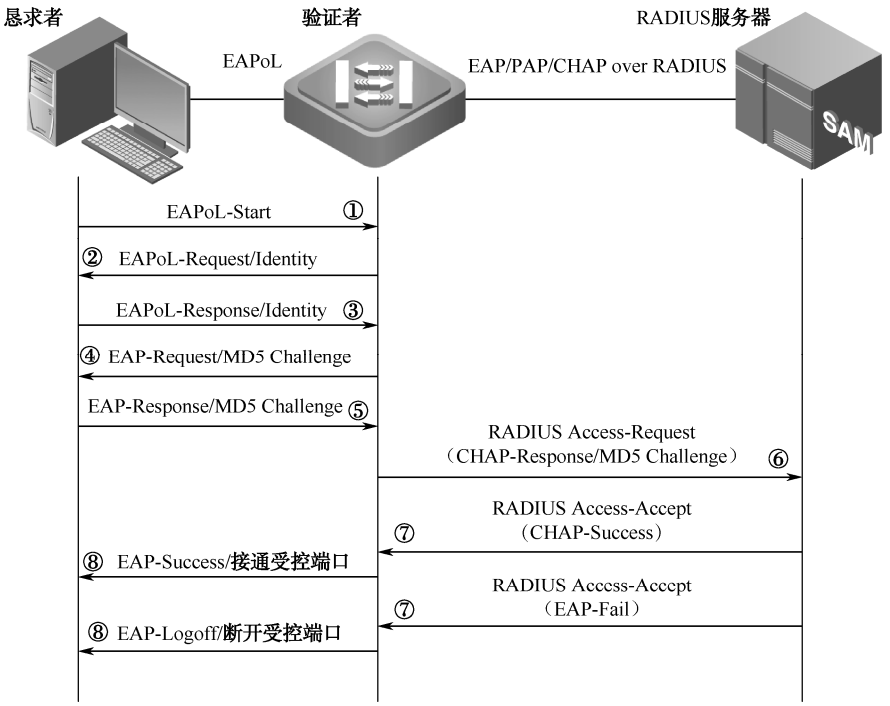


图 6.5　EAP 终结模式验证过程

⑥ 交换机会将客户端的用户名、MD5 挑战值和客户端计算的散列值一同发送给 RADIUS 服务器。

⑦ RADIUS 服务器通过将收到的客户端的散列值与自己计算的散列值进行比较，如果相同则表示用户合法，验证通过，并返回 RADIUS Accept 报文，其中包含 EAP-Success 信息。

⑧ 交换机收到验证通过的信息后，将连接客户端的接口“开放”，并发送 EAP-Success 报文给客户端，以通过客户端验证通过。或者没有验证失败，交换机受控接口仍处于接口断开状态。

从上述工作过程可以看出，在 EAP 终结模式中，MD5 挑战值是由交换机生成的，随后交换机会将客户端的用户名、MD5 挑战值和客户端计算的散列值一同发送给 RADIUS 服务器，再由 RADIUS 服务器进行验证。对于 EAP 终结模式，交换机与 RADIUS 服务器之间只交换两条消息，减少了他们之间的信息交互量，减轻了 RADIUS 服务器的负担。

4. 802.1X 基本配置

（1）AAA 与 RADIUS 配置。

第 1 步：启用 AAA 功能。由于 802.1X 是基于 AAA 工作的，所以在启用 802.1X 之前需要在交换机上先启用 AAA 功能。启用 AAA 功能的命令格式如下：

```
                          启用AAA安全模型
                                │
ruijie(config)# aaa new-model
       │
交换机名称为ruijie
处于全局配置模式
```

第 2 步：配置 RADIUS 服务器 IP 地址。交换机需要将 802.1X 验证请求报文发送至特定的 RADIUS 服务器以验证用户的身份，指定特定 RADIUS 服务器地址的命令格式如下：

配置RADIUS服务器地址命令　RADIUS服务器IP地址　指定记账UDP接口号

ruijie(config)# radius-server host *{hostname | ip-address}* **[auth-port** *port-number*] **[acct-port** *port-number*]

设备名称为ruijie 处于全局配置模式　RADIUS服务器DNS名称　指定验证UDP接口号

第 3 步：设置交换机与 RADIUS 服务器通信密码。为了保证交换机与 RADIUS 服务器之间通信的安全性，需要提供共享密码。设置共享密码的命令格式如下：

配置RADIUS服务器验证密码命令

ruijie(config)# radius-server key *text-string*

设备名称为ruijie 处于全局配置模式　密码值

第 4 步：定义 802.1X 验证方法列表。使用如下命令定义 802.1X 验证方法列表：

定义验证的用户行为　定义验证的方法

ruijie(config)# aaa authentication dot1x {default | *list-name***}** *method1 [method2...]*

设备名称为ruijie 处于全局配置模式　定义方法列表名称

以上命令在前面章节中已经详细介绍过，在此不再详细描述。

（2）启用 802.1X。配置完 RADIUS 服务器相关参数后，即完成了 802.1X 的准备工作。剩下的工作就是启用接口的 802.1X 验证功能。启用了 802.1X 验证功能的接口则成为受控接口，连接在该接口下的用户要通过验证才能访问网络，然而，在非受控接口下的用户可以直接访问网络。

使用如下命令启用接口的 802.1X 验证功能：

启用端口802.1X验证功能

ruijie(config-if)# dot1x port-control auto

设备名称为ruijie 处于接口配置模式

当接口配置了此命令后，接口将只接收 EAPoL 报文，并且只有验证通过后，接口才会接收其他报文。

默认情况下，接口不参加 802.1X 验证。使用该命令的 no 选项，恢复接口的 802.1X 默认功能，如 ruijie(config-if)#no dot1x port-control。

例如：将交换机 Fastethernet 0/1 接口设置为受控接口。

```
ruijie(config)#interface fa 0/1
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#end
```

例如：将交换机 Fastethernet 0/1 接口恢复为非受控接口。

```
ruijie(config)#interface fa 0/1
ruijie(config-if)#no dot1x port-control
ruijie(config-if)#exit
```

（3）应用 802.1X 验证方法列表。启用 802.1X 验证功能后，默认情况下，交换机将使用默认的验证列表（default）进行验证。可以在全局配置模式下使用如下命令关联 802.1X 验证方法列表：

定义的方法列表名称

ruijie(config)#dot1x authentication {default | *list-name*}

设备名称为ruijie
处于全局配置模式

其中，*list-name* 为定义的 802.1X 验证方法列表名称。

如果 AAA 打开，则登录验证使用默认验证方法列表。使用该命令的 no 选项，将删除已设置的登录认证方法列表，如 ruijie(config)# no dot1x authentication {default | list-name}。

例如：在交换机 Fa0/1 及 Fa0/10 接口上实施 802.1X 验证功能。

```
ruijie(config)#aaa new-model
ruijie(config)#radius-server host 192.168.217.64
ruijie(config)#radius-server key key123
ruijie(config)#aaa authentication dot1x authen group radius
ruijie(config)#dot1x authentication authen
ruijie(config)#interface fa 0/1
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#exit
ruijie(config)#interface fa 0/10
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#end
ruijie#
```

例如：将交换机 Fa0/10 接口恢复为非受控接口。

```
ruijie(config)#interface fa 0/10
ruijie(config-if)#no dot1x port-control
ruijie(config-if)#end
```

（4）配置 802.1X 验证方式。在 802.1X 标准中规定的验证方式为 EAP-MD5，除此之外还提供了 PAP 和 CHAP 验证方式。CHAP 比 PAP 安全性高些，PAP 采用明文传递用户名和密码，CHAP 采用密文传递密码信息。

当使用 PAP 或 CHAP 验证方式时，交换机执行 EAP 终结模式，即 EAP 协议在交换机侧终结，所以 RADIUS 服务器无须支持 EAP 属性。如前所述，对于 EAP 终结模式，交换机与 RADIUS 服务器之间只交换两条信息，减少了信息交互量，减轻了 RADIUS 服务器负担。

配置交换机 802.1X 验证方式的命令格式如下：

设置dot1x验证方式命令　　CHAP或PAP方式验证

ruijie(config)# dot1x auth-mode *{eap-md5 | chap | pap}*

交换机名称为ruijie
处于全局配置模式　　EAP-MD5 方式验证

其中，*eap-md5* 为 802.1X 采用 EAP-MD5 方式验证；*chap* 为 802.1X 采用 CHAP 方式验证；*pap* 为 802.1X 采用 PAP 方式验证；默认验证方式为 EAP-MD5。

默认情况下采用 EAP-MD5 验证方式，使用该命令的 no 选项可以恢复默认设置方式，如 ruijie(config)# no dot1x auth-mode。

例如：设置交换机的 802.1X 认证方式为 CHAP。

```
ruijie(config)#dot1x auth-mode chap
ruijie(config)#end
```

例如：将交换机的 802.1X 验证方式恢复为默认方式 EAP-MD5。

```
ruijie(config)#no dot1x auth-mode
```

（5）配置 EAPoL 帧携带 TAG。IEEE 802.1X 标准中规定 EAPoL 报文不能携带 VLAN TAG 标记。但是某些应用环境下，需要 EAPoL 帧携带 VLAN TAG 标记，以使 EAPoL 帧能够被发送至相应的 VLAN 中。典型应用是在汇聚层交换机启用 802.1X 验证功能。

使用如下命令启用 EAPoL 帧携带 VLAN TAG 功能：

默认情况下，EAPoL 帧不携带 VLAN TAG 标记。使用该命令的 no 选项恢复为默认设置值，如 ruijie(config)# no dot1x eapol-tag。

例如：设置 802.1X 帧带 TAG。

```
ruijie(config)#dot1x eapol-tag
ruijie(config)#end
```

（6）802.1X 监控与维护。

查看 802.1X 的运行状态及配置信息：

```
ruijie#show dot1x
```

查看 802.1X 定时器信息：

```
ruijie#show dot1x timeout quiet-period
ruijie#show dot1x timeout re-authperiod
ruijie#show dot1x timeout server-timeout
ruijie#show dot1x timeout supp-timeout
ruijie#show dot1x timeout tx-period
```

查看 802.1X 计数器信息：

```
ruijie#show dot1x reauth-max
ruijie#show dot1x max-req
```

查看交换机接口控制信息：

```
ruijie#show dot1x port-control [interface interface]
```

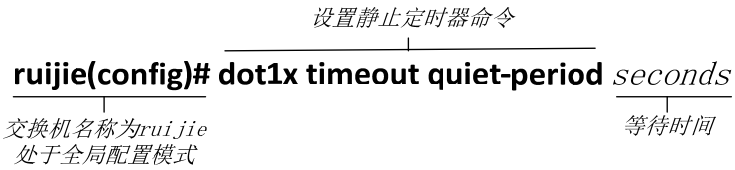
6.3　扩展知识

802.1X 使用多个定时器来保证验证过程的正常进行及验证组件间的故障检测等，下面分别介绍各定时器的功能。

1. 静止定时器（quiet-period）

当用户验证失败时，设备将等待一段时间后，才允许用户再次验证。Quiet Period 的时间长度便是允许再验证的时间间隔。该值的作用是避免设备受恶意攻击。Quiet Period 的默认间隔为 10 秒，我们可以通过设定较短的 Quiet Period 值，使用户可以更快地进行再验证。

设置静止定时器的命令格式如下：



其中，*seconds* 为设备验证失败后到允许尝试重新验证的等待时间。值的范围为 0～65535，单位为秒，默认值为 10 秒。使用该命令的 no 选项将该设置恢复为默认值，如 ruijie(config)#no dot1x timeout quiet-period。

例如：设置重新验证等待时间为 1000 秒。

```
ruijie(config)#dot1x timeout quiet-period 1000
ruijie(config)#end
```

例如：将重新验证等待时间恢复为默认值 10 秒。

```
ruijie(config)#no dot1x timeout quiet-period
```

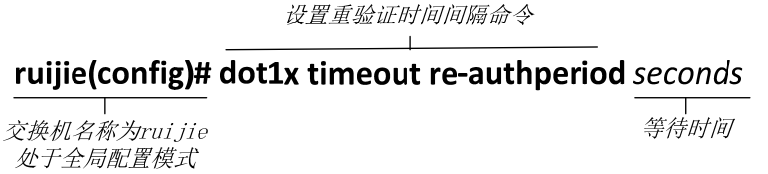
2. 重验证定时器（re-authperiod）

802.1X 能定时主动要求用户重新验证，这样可以防止已通过验证的用户不会被其他用户冒用，还可以检测用户是否断线，使记费更准确。除了可以设定重验证的开关，我们还可以定义重验证的间隔。

打开定时重验证功能的命令格式如下：

默认情况下不要求恳请者定期重新认证。使用该命令的 no 选项将设置恢复为默认值，如 ruijie(config)#no dot1x re-authentication。

启动重验证功能后，系统每到默认重验证时间 3600 秒就要求客户端进行重新验证身份。重验证时间间隔也可以通过如下命令设置：



其中，*seconds* 为验证周期，值的范围为 0～65535，单位为秒，默认值为 3600 秒。使用该命令的 no 选项将该设置恢复为默认值，如 ruijie(config)# no dot1x timeout re-authperiod。

例如：设置重认证周期为 1000 秒。

```
ruijie(config)#dot1x timeout re-authperiod 1000
ruijie(config)#end
```

例如：将从验证时间恢复为默认值。

```
ruijie(config)#no dot1x timeout re-authperiod
```

当交换机进行重验证时，但到达重验证间隔时间后还没有收到客户端的响应时，交换机将重新发送验证请求，设置可以发送的重验证次数的命令格式如下：

其中，*count* 为最大重验证次数，默认值为 3。使用该命令的 no 选项将该设置恢复成默认值，如 ruijie(config)# no dot1x reauth-max。

例如：设置重验证最大次数为 5。

```
ruijie(config)#dot1x reauth-max 5
ruijie(config)#end
```

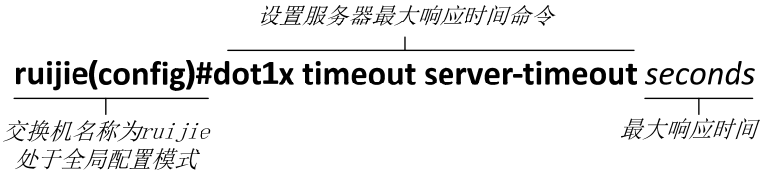
例如：将重验证最大次数恢复为默认值 3。

```
ruijie(config)#no dot1x reauth-max
```

3. 服务器超时定时器（server-timeout）

服务器超时定时器的值指的是 Radius Server 的最大响应时间，若在该时间内，设备没有收到 Radius Server 的响应，将认为本次验证失败。

设置服务器超时定时器的命令格式如下：



其中，*seconds* 为设备和验证服务器之间验证交互的超时时间，值的范围为 0～65535 秒，默认值 5 秒。使用该命令的 no 选项将该设置恢复为默认值，如 ruijie(config)# no dot1x timeout server-timeout。

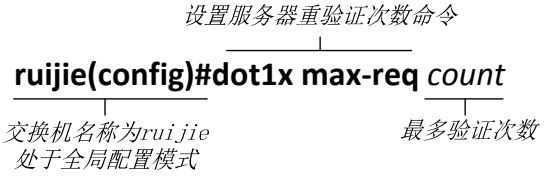
例如：设置服务器交互超时时间为 10 秒。

```
ruijie(config)#dot1x timeout server-timeout 10
ruijie(config)#end
```

例如：将服务器的最大响应时间恢复为默认值 5 秒。

```
ruijie(config)#no dot1x timeout server-timeout
```

在 dot1x 设备和服务器交互过程中，如果 dot1x 设备在一定时间内没有收到服务器的响应，则 dot1x 设备将再次向服务器发起请求。使用下述命令设置向服务器请求允许的最大次数。



其中，*count* 为允许的最大验证请求的次数，默认值为 3。使用该命令的 no 选项将该设置恢复成默认值，如 ruijie(config)# no dot1x max-req。当重发验证请求次数达到设定值后，交换机将认为服务器不可用。

例如：设置 802.1X 设备与验证服务器之间的最大验证重传 7 次，超过 7 次后认为服务器不可用。

```
ruijie(config)#dot1x max-req 7
```

例如：将 802.1X 设备的最大验证重传次数恢复为默认值 3。

```
ruijie(config)#no dot1x max-req
```

4. 客户端超时定时器（supp-timeout）

客户端超时定时器的值表示当交换机向客户端发送 EAP-Request/MD5Challenge 报文请求 MD5 散列值后，等待客户端最大响应时间。

设置客户端超时定时器的命令格式如下：

设置等待客户端最大响应时间命令

ruijie(config)# dot1x timeout supp-timeout *seconds*

交换机名称为ruijie
处于全局配置模式

超时时间

其中，*seconds* 为设备和恳请者之间验证交互的超时时间。值的范围为 0～65535 秒，默认值为 3 秒。使用该命令的 no 选项将该设置恢复为默认值，如 ruijie(config)#no dot1x timeout supp-timeout。

例如：设置客户端交互超时时间为 10 秒。

```
ruijie(config)#dot1x timeout supp-timeout 10
ruijie(config)#end
```

例如：将交换机等待客户端最大响应时间恢复为默认值 3 秒。

```
ruijie(config)#no dot1x timeout supp-timeout
```

5. 发送超时定时器（tx-period）

当客户端发起验证过程后，交换机需要向客户端发送 EAP-Request/Identity 报文，请求客户端发送用户名信息。当交换机超过了 tx-perriod 定时器设定的时间间隔后还没有收到客户端的 EAP-Response/Identity 响应报文时，交换机则重发请求报文。

设置发送超时定时器的命令格式如下：

设置发送超时定时器时间命令

ruijie(config)# dot1x timeout tx-period *seconds*

交换机名称为ruijie
处于全局配置模式

超时时间

其中，*seconds* 为重传周期。值的范围为 0～65535 秒，默认值为 3 秒。使用该命令的 no 选项将该设置恢复为默认值，如 ruijie(config)# no dot1x timeout tx-period。

例如：配置定时器的实例如下。

```
ruijie(config)#dot1x timeout quiet-period 5
ruijie(config)#dot1x timeout server-timeout 3
ruijie(config)#dot1x timeout supp-timeout 2
ruijie(config)#dot1x timeout tx-period 5
```

6.4 演示实例

1. 背景描述

接入层交换机接口启用 802.1X 验证，实现对接入网络的用户进行控制，该方案在实际应用中被普遍采用，如图 6.6 所示。

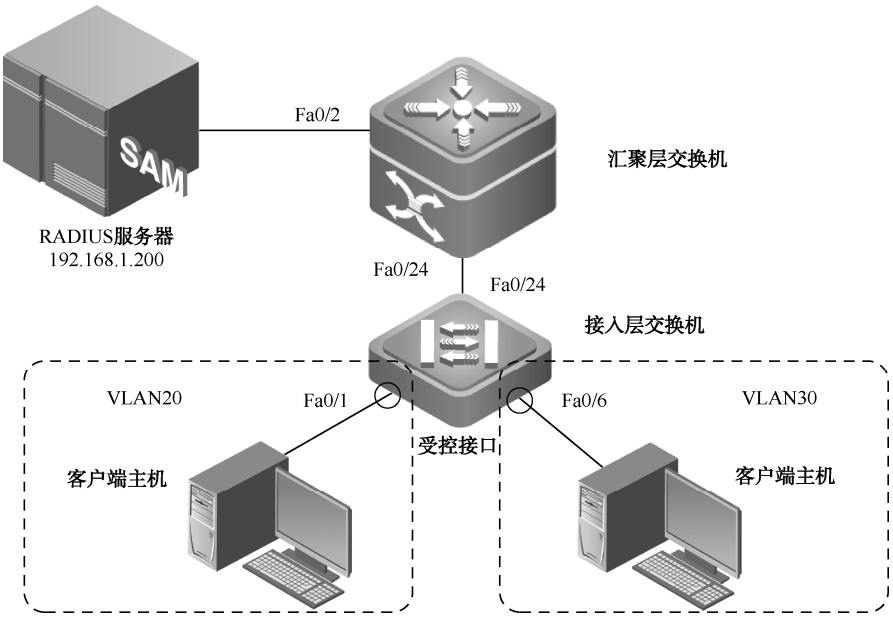


图 6.6　接入层接口启用 802.1X

具体要求如下：

（1）用户支持 802.1X，即要装有 802.1X 的客户端软件（Windows XP 自带，Star-supplicant 或其他符合 IEEE 802.1X 标准的客户端软件）；

（2）接入层设备支持 IEEE 802.1X；

（3）有一台（或多台）支持标准 RADIUS 的服务器作为验证服务器。

该方案的配置要点如下：

（1）与 RADIUS Server 相连的接口及上联接口，配置成非受控接口，以便设备能正常地与服务器进行通信，以及使已验证用户能通过上联接口访问网络资源；

（2）与用户连接的接口要设置为受控接口，以实现对接入用户的控制，用户必须通过验证才能访问网络资源。

该方案的特点如下：

（1）每台支持 802.1X 的设备所负责的客户端少，验证速度快。各台设备之间相互独立，设备的重启等操作不会影响到其他设备所连接的用户；

（2）用户的管理集中于 RADIUS Server 上，管理员不必考虑用户连接在哪台设备上，便于管理员的管理；

（3）管理员可以通过网络管理接入层的设备。

2. 操作步骤

接入层交换机配置代码如下：

```
ruijie(config)# VLAN 20
ruijie(config-VLAN)#exit
ruijie(config)# VLAN 30
ruijie(config- VLAN)#exit
ruijie(config)#interface fastethernet 0/1
```

```
ruijie(config-if)#switchport access VLAN 20
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/6
ruijie(config-if)#switchport access VLAN 30
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#switchport mode trunk
ruijie(config-if)#exit
ruijie(config)#aaa new-model
ruijie(config)#aaa authentication dot1x test group radius
ruijie(config)#dot1x authentication test
ruijie(config)#radius-server host 192.168.1.200
ruijie(config)#radius-server key key123
ruijie(config)#interface fastethernet 0/1
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/6
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#end
ruijie#
```

汇聚层交换机配置代码如下：

```
ruijie(config)#VLAN 10
ruijie(config-VLAN)#exit
ruijie(config)# VLAN 20
ruijie(config- VLAN)#exit
ruijie(config)# VLAN 30
ruijie(config- VLAN)#exit
ruijie(config)#interface fastethernet 0/2
ruijie(config-if)#switchport access vlan 10
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#switchport mode trunk
ruijie(config-if)#exit
ruijie(config)#interface VLAN 10
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config)#interface VLAN 20
ruijie(config-if)#ip address 192.168.2.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config)#interface VLAN 30
ruijie(config-if)#ip address 192.168.3.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#end
ruijie#
```

6.5　训练任务

1. 背景描述

在网络工程中，有时接入层交换机不支持 802.1X，此时需要在汇聚层启用 802.1X 验证，实现对用户的验证，如图 6.7 所示。

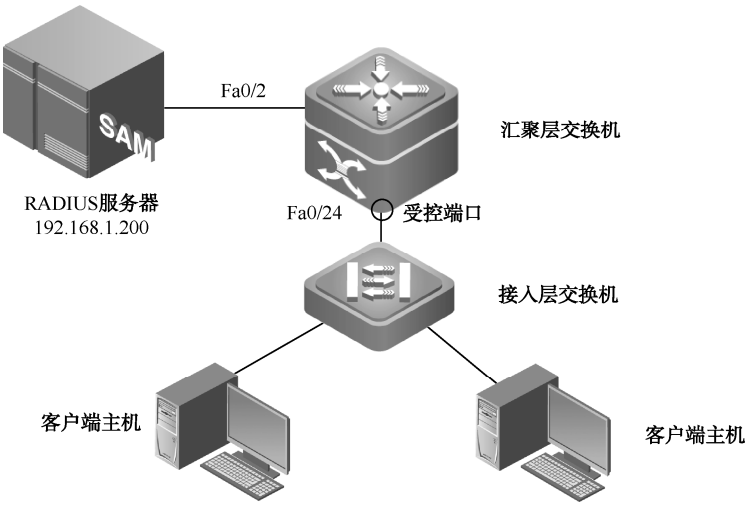


图 6.7　汇聚层接口启用 802.1X

具体要求如下：

（1）用户支持 802.1X，即要装有 802.1X 的客户端软件（Windows XP 自带，Star-supplicant 或其他符合 IEEE 802.1X 标准的客户端软件）；

（2）接入层设备支持要能透传 IEEE 802.1X 帧(EAPoL)；

（3）汇聚层设备支持 802.1X（扮演验证者角色）；

（4）有一台（或多台）支持标准 RADIUS 的服务器作为验证服务器。

该方案的配置要点：

（1）与 RADIUS Server 相连的接口及上联接口，配置成非受控接口，以便设备能正常地与服务器进行通信，以及使已验证用户能通过上联接口访问网络资源；

（2）与接入层设备连接的接口要设置为受控接口，以实现对接入用户的控制，用户必须通过验证才能访问网络资源。

该方案的特点：

（1）由于是汇接层设备，网络规模大，下接用户数多，对设备的要求高，若该层设备发生故障，将导致大量用户不能正常访问网络；

（2）用户的管理集中于 RADIUS Server 上，管理员无须考虑用户连接在哪台设备上，便于管理员的管理；

（3）接入层设备可以使用较廉价的非网管型设备（只要支持 EAPoL 帧透传即可）；

（4）管理员不能通过网络直接管理接入层设备。

2. 操作提示

汇聚层交换机配置代码如下：

```
ruijie(config)#VLAN 10
ruijie(config-VLAN)#exit
ruijie(config)#interface fastethernet 0/2
ruijie(config-if)#switchport access VLAN 10
ruijie(config-if)#exit
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#switchport mode trunk
ruijie(config-if)#exit
ruijie(config)#interface VLAN 10
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
ruijie(config-if)#no shutdown
ruijie(config-if)#exit
ruijie(config)#aaa new-model
ruijie(config)#aaa authentication dot1x test group radius
ruijie(config)#dot1x authentication test
ruijie(config)#radius-server host 192.168.1.200
ruijie(config)#radius-server key key123
ruijie(config)#interface fastethernet 0/24
ruijie(config-if)#dot1x port-control auto
ruijie(config-if)#exit
ruijie(config)#dot1x eapol-tag
ruijie(config)#end
ruijie#
```

练习题

1. 选择题

（1）IEEE 802.1X（Port-Based Network Access Control）是一个基于（　　）的网络访问控制标准。

A．接口　　B．命令行　　C．服务　　D．服务器

（2）下列不是 IEEE 802.1X 标准验证体系中的组成部分的是（　　）。

A．恳请者　　B．验证者　　C．验证服务器　　D．接入集线器

（3）下面说法错误的是（　　）。

A．在客户端与交换机之间，EAP 报文直接被封装到 LAN 协议中，形成 EAPoL 报文

B．在交换机与 RADIUS 服务器之间，EAP 协议报文被封装到 RADIUS 报文中

C．在交换机与 RADIUS 服务器之间还可以使用 RADIUS 协议交互 PAP 和 CHAP 报文

D．交换机在整个验证过程中参与验证，验证通过后才打开交换机接口

2. 简答题

（1）描述 802.1X 验证体系构成。

（2）说明 802.1X 两种验证模式（EAP 中继模式和 EAP 终结模式）的区别。

（3）说明配置 802.1X 的主要步骤有哪些。

项目 7　RIP 路由协议与高级配置

在网络互联技术基础篇中学习了 RIP 路由协议的基本功能及配置，本章内容将在前面学习的基础上，进一步了解 RIP 路由协议的更多特性及配置技能，包括 RIP 路由汇总、RIP 验证、时钟调整等。

知识点、技能点

1. 掌握 RIP 协议路由汇总功能及配置技能;
2. 掌握 RIP 验证功能及配置技能;
3. 掌握时钟配置技能。

7.1　问题提出

某公司网络采用 RIP 路由协议进行配置，但一方面由于子网数量多，且网络不够稳定，造成公司三层设备路由表波动大，设备负担重；另一方面，公司出于安全考虑，希望能够配置 RIP 验证，实现只有合法设备才能够学习到路由信息。请您根据公司要求制定合适网络实施方案，并完成相应配置工作。

7.2　相关知识

1. 回顾 RIP 路由协议的基本特性及应用配置

通过前面的学习我们已经了解 RIP 的基本特性，下面我们再简单地回顾一下相关内容。

RIP（Routing Information Protocol）路由协议是一种相对古老，在小型及铜介质网络中得到了广泛应用的路由协议。RIP 采用距离向量算法，是一种距离向量协议。目前，RIP 有 RIPv1 和 RIPv2 两个版本。RIP 使用 UDP 报文交换路由信息，UDP 接口号为 520。通常情况下 RIPv1 报文为广播报文；而 RIPv2 报文为组播报文，组播地址为 224.0.0.9。

RIP 每隔 30 秒向外发送一次更新报文。如果设备经过 180 秒没有收到来自对端的路由更新报文则将所有来自此设备的路由信息标志为不可达，路由进入不可达状态后，120 秒内仍未收到更新报文就将这些路由信息从路由表中删除。

RIP 使用跳数来衡量到达目的地的距离，称为路由量度。在 RIP 中，设备到与它直接相连网络的跳数为 0；通过一个设备可达的网络的跳数为 1，其余依此类推；不可达网络的跳数为 16。

（1）创建 RIP 路由进程。设备要运行 RIP 路由协议，首先需要创建 RIP 路由进程，并定义与 RIP 路由进程关联的网络。

要创建 RIP 路由进程，在全局配置模式中执行以下命令：

```
ruijie(config)#router rip
ruijie(config-router)#network network-number wildcard
```

其中，*network-number* 为直连网络的网络号。该网络号为自然类网络号，IP 地址属于该自然网络

的所有接口都可发送和接收 RIP 数据包。

wildcard 为定义 IP 地址比较比特位，0 表示精确匹配，1 表不做比较。

用户可以同时配置 network-number 和 wildcard 参数，使落在该地址范围内的接口地址参与 RIP 运行。如果未配置 wildcard 参数，将按照默认有类地址范围来处理，使落在该有类地址范围内的接口地址参与 RIP 运行。

只有接口地址落在 RIP 定义的网络列表中，该接口才可以对外发送 RIP 路由更新报文，并接收 RIP 路由更新报文。

在下面的配置代码中，定义了与 RIP 关联的两个网络号，允许落在 192.168.12.0/24 和 172.16.0.0/24 范围内的接口地址参与 RIP 运行：

```
ruijie(config)#router rip
ruijie(config-router)#network 192.168.12.0
ruijie(config-router)#network 172.16.0.0 0.0.0.255
```

（2）定义 RIP 版本。RIP 有 RIPv1 和 RIPv2 两个版本。RIPv1 为有类路由、不支持 VLSM、不支持不连续子网、广播方式更新、自动汇总（不可关闭）、不支持手动汇总、不支持路由验证；RIPv2 为无类路由、支持 VLSM、支持不连续子网、组播方式更新（224.0.0.9）、自动汇总（可关闭）、支持手动汇总、支持路由验证。

默认情况下，可以接收 RIPv1 和 RIPv2 的数据包，但是只发送 RIPv1 的数据包。可以通过配置，只接收和发送 RIPv1 的数据包，也可以只接收和发送 RIPv2 的数据包。

要配置设备只接收和发送指定版本的数据包，在路由进程配置模式中执行以下命令：

```
ruijie(config-router)#version {1 | 2}
```

以上命令使设备默认情况下只接收和发送指定版本的数据包，如果需要可以更改每个接口的默认操作行为。

要配置接口只发送哪个版本的数据包，在接口配置模式中执行以下命令。

指定只发送 RIPv1 数据包的命令如下：

```
ruijie(config-if)#ip rip send version 1
```

指定只发送 RIPv2 数据包的命令如下：

```
ruijie(config-if)#ip rip send version 2
```

指定只发送 RIPv1 和 RIPv2 数据包的命令如下：

```
ruijie(config-if)#ip rip send version 1 2
```

要配置接口只接收哪个版本的数据包，在接口配置模式中执行以下命令。

指定只接收 RIPv1 数据包的命令如下：

```
ruijie(config-if)#ip rip receive version 1
```

指定只接收 RIPv2 数据包的命令如下：

```
ruijie(config-if)#ip rip receive version 2
```

指定只接收 RIPv1 和 RIPv2 数据包的命令如下：

```
ruijie(config-if)#ip rip receive version 1 2
```

由于 RIPv1 和 RIPv2 工作方式及报文格式不同，如果网络中路由器的 RIP 版本不同，就会出现兼容性问题，致使运行 RIPv1 的路由器和 RIPv2 的路由器之间不能学习到相关路由信息。此时就需要使用上述命令定义接口发送或接收的 RIPv1、RIPv2 或 RIPv1 和 RIPv2。

如图 7.1 所示，路由器 R1 运行 RIPv1 路由协议，路由器 R2 运行 RIPv2 路由协议，R1 无法学习到 R2 的路由信息，同样 R2 也学习不到 R1 的路由信息。为了解决 RIP 协议兼容性问题，在路由器 R2

的 Fa0/0 接口配置接收、发送 RIPv1 和 RIPv2 的数据包，这样配置后，R1 能够学习到 R2 的路由信息，同样 R2 也能够学习到 R1 的路由信息。

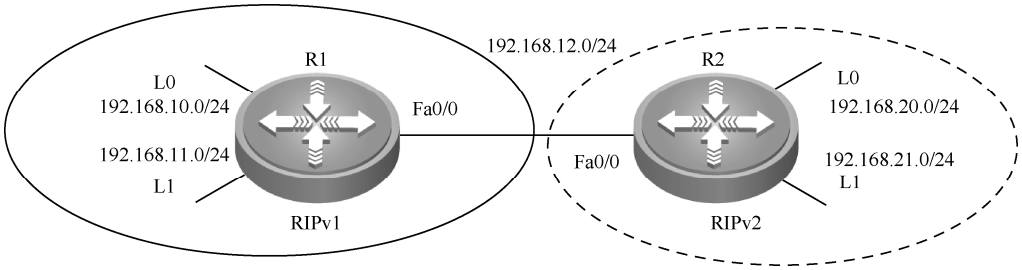


图 7.1　RIPv1 和 RIPv2 兼容性

路由器 R1 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R1
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 0
R1(config-if)#ip address 192.168.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 1
R1(config-if)#ip address 192.168.11.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#router rip
R1(config-router)#version 1
R1(config-router)#network 192.168.10.0
R1(config-router)#network 192.168.11.0
R1(config-router)#network 192.168.12.0
R1(config-router)#end
R1#
```

路由器 R2 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R2
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#ip rip send version 1 2
R2(config-if)#ip rip receive version 1 2
R2(config-if)#exit
R2(config)#interface loopback 0
```

```
R2(config-if)#ip address 192.168.20.1 255.255.255.0
R2config-if)#no shutdown
R2(config-if)#exit
R2(config)#interface loopback 1
R2(config-if)#ip address 192.168.21.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#network 192.168.12.0
R2(config-router)#network 192.168.20.0
R2(config-router)#network 192.168.21.0
R2(config-router)#end
R2#
```

使用 show ip route 命令查看路由器 R1、R2 的路由表中是否学习到相应的路由信息。

（3）水平分割配置。多台设备连接在 IP 广播类型网络上，又运行距离向量路由协议时，就有必要采用水平分割的机制以避免路由环路的形成。水平分割可以防止设备将某些路由信息从学习到这些路由信息的接口通告出去，这种行为优化了多台设备之间的路由信息交换。

然而对于非广播多路访问网络（如帧中继、X.25 网络），水平分割可能造成部分设备学习不到全部的路由信息。在这种情况下，可能需要关闭水平分割。如果一个接口配置了次 IP 地址，也需要注意水平分割的问题。

要配置关闭或打开水平分割，在接口配置模式中执行以下命令。

关闭水平分割命令：

```
ruijie(config-if)#no ip split-horizon
```

打开水平分割命令：

```
ruijie(config-if)#ip split-horizon
```

所有接口的默认行为设置为开启水平分割。

（4）RIP 路由自动汇总。RIP 路由自动汇总，就是当子网路由穿越有类网络边界时，将自动汇总成有类网络路由。RIPv2 默认情况下将进行路由自动汇总，RIPv1 不支持该功能。

RIPv2 路由自动汇总的功能，提高了网络的伸缩性和有效性。如果有汇总路由存在，在路由表中将看不到包含在汇总路由内的子路由，这样可以大大缩小路由表的规模，但也容易造成路由黑洞。

有时可能希望学到具体的子网路由，而不愿意只看到汇总后的网络路由，这时需要关闭路由自动汇聚功能。默认情况下启用路由自动汇总功能。

要配置路由自动汇总，在 RIP 路由进程模式中执行以下命令。

关闭路由自动汇总功能：

```
ruijie(config-router)#no auto-summary
```

打开路由自动汇总功能：

```
ruijie(config-router)#auto-summary
```

如图 7.2 所示，路由器 R1 和 R2 均执行 RIPv2 协议，通过关闭自动汇总功能获得子网路由信息。

图 7.2　RIP 自动汇总

路由器 R1 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R1
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 0
R1(config-if)#ip address 172.16.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 1
R1(config-if)#ip address 172.16.11.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 2
R1(config-if)#ip address 10.10.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#router rip
R1(config-router)#version 2
R1(config-router)#network 172.16.10.0
R1(config-router)#network 172.16.11.0
R1(config-router)#network 10.10.10.0
R1(config-router)#network 192.168.12.0
R1(config-router)#no auto-summary
R1(config-router)#end
R1#
```

路由器 R2 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R2
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#interface loopback 0
```

```
R2(config-if)#ip address 192.168.20.1 255.255.255.0
R2config-if)#no shutdown
R2(config-if)#exit
R2(config)#interface loopback 1
R2(config-if)#ip address 192.168.21.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#network 192.168.12.0
R2(config-router)#network 192.168.20.0
R2(config-router)#network 192.168.21.0
R2(config-router)#no auto-summary
R2(config-router)#end
R2#
```

使用 show ip route 命令查看路由器 R1、R2 的路由表中是否学习到相应的路由信息。然后，使用 auto-summaty 命令启用自动汇总功能，再查看路由器 R1、R2 的路由表，并对比关闭与启用自动汇总功能对路由表的影响。

2. RIP 路由协议的高级特性及应用配置

1）RIP 手动汇总

只有 RIPv2 才支持手动汇总功能，即将本地的同一个主网的多个子网通过命令汇总为一条路由通告出去，这样可以大大减小对方路由表的大小，从而大大优化网络的目的。

为了启用 RIP 手动汇总功能，在接口配置模式下使用命令 ip summary-address rip，命令格式如下：

启用RIP手动汇总功能　　*汇总子网掩码*

ruijie(config-if)#ip summary-address rip *ip-address ip-network-mask*

设备名称为ruijie 处于接口配置模式　　*汇总地址*

其中，*ip-address* 为汇总网络地址，也即汇总后的网络地址；*ip-network-mask* 为指定 IP 地址进行路由汇总的子网掩码。

默认情况下，RIP 自动汇聚到有类网络边界。使用该命令的 no 形式关闭指定地址或子网的汇聚，如 ruijie(config-if)#no ip summary-address rip *ip-address ip-network-mask*。

例如：在某网络配置中，关闭了 RIPv2 的路由自动汇总。配置接口汇总，接口 FastEthernet 1/0 将通告汇总后的路由 172.16.0.0/16。

```
ruijie(config)#interface FastEthernet 1/0
ruijie(config-if)#ip summary-address rip 172.16.0.0 255.255.0.0
ruijie(config-if)#ip address    172.16.1.1 255.255.255.0
ruijie(config)#router rip
ruijie(config-router)#network 172.16.0.0
ruijie(config-router)#version 2
ruijie(config-router)#no auto-summary
```

例如：对于前面图 7.2 所示网络中 R1 的 Fa0/0 接口进行手动路由汇总，实现将 172.16.10.0/24 及 172.16.11.0/24 汇总成一条 172.16.0.0/16 路由通告路由器 R2。

实现手动路由汇总的路由器 R1 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R1
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#ip summary-address rip 172.16.0.0 255.255.0.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 0
R1(config-if)#ip address 172.16.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 1
R1(config-if)#ip address 172.16.11.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback2
R1(config-if)#ip address 10.10.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#router rip
R1(config-router)#version 2
R1(config-router)#network 172.16.10.0
R1(config-router)#network 172.16.11.0
R1(config-router)#network 10.10.10.0
R1(config-router)#network 192.168.12.0
R1(config-router)#no auto-summay
R1(config-router)#end
R1#
```

通过查看路由器 R2 的路由表，检查路由器 R1 的手动汇总效果。

2）配置 RIP 验证

配置 RIP 验证可以大大增加 RIP 路由信息更新的安全性，只有双方通过验证的情况下才可以相互通告 RIP 路由更新报文。RIPv1 不支持验证，如果设备配置 RIPv2 路由协议，可以在相应的接口配置验证。 RIPv2 支持两种验证方式：明文验证和 MD5 验证。默认的验证方式为明文验证。

配置 RIP 验证的步骤如下。

第 1 步：定义 RIP 验证的方式。要定义 RIP 验证的模式，请使用接口配置命令 ip rip authenticationmode，命令格式如下：

其中，*text* 为 RIP 验证模式为明文验证；*md5* 为 RIP 验证模式为 MD5 验证。默认验证模式为明文验证，该命令的 no 形式恢复默认 RIP 验证模式，如 ruijie(config-if)# no ip rip authentication mode。

提示：当配置 RIP 验证时，所有需要直接交换 RIP 路由信息的设备，所配置的 RIP 验证模式必须一致，否则 RIP 数据包交换将失败。

例如：在路由器 Serial 0/0 接口上配置 RIP 认证模式为 MD5。

```
ruijie(config)#interface serial 0/0
ruijie(config-if)#ip rip authentication mode md5
```

第 2 步：定义验证密码。定义验证密码分为明文及密码方式，下面分别进行描述。

（1）明文验证方式

明文验证就是双方发起验证时验证密码以明文方式发送给对方。如果要使用明文验证，用户可以直接使用 ip rip authentication text-password 方式配置明文验证字符串，也可以通过关联密钥串获取明文验证字符串，后者优先级高于前者。

方法一：直接配置明文验证密码。

设置 RIP 明文验证密码的命令格式如下：

设置RIP明文验证密码命令

ruijie(config-if)#ip rip authentication text-password *password-string*

设备名称为ruijie
处于接口配置模式

验证字符串

其中，*password-string* 为明文验证的字符串，长度为 1～16 字节。默认情况下，未设置明文验证字符串，使用该命令的 no 形式删除明文验证的字符串，如 ruijie(config-if)#no ip rip authentication text-password。

例如：在路由器 Serial 0/0 接口上启用 RIP 明文验证，并设置明文验证字符串为 ruijie。

```
ruijie(config)#interface serial 0/0
ruijie(config-if)#ip rip authentication text-password ruijie
```

方法二：关联密钥串获取明文验证字符串。

在使用关联密钥串之前，需要先定义密钥链、定义密钥及密钥字符串。下面分别进行描述。

① 定义密钥链。在全局配置模式下使用 key chain 命令定义密钥链，并进入密钥链配置模式。要使密钥链生效，必须配置至少一个密钥。

提示：如果密钥链在接口配置中指定了，但是没有利用 key chain 全局配置命令进行密钥串定义，将不进行 RIP 数据包验证。

定义密钥链的命令格式如下：

定义密钥链命令

ruijie(config)#key chain *key-chain-name*

设备名称为ruijie
处于全局配置模式

密钥链名称

其中，*key-chain-name* 为密钥链名称。默认情况下，没有定义任何密钥链。使用该命令的 no 形式删除指定密钥链的定义，如 ruijie(config)#no key chain *key123*，其中 *key123* 为要删除的密钥链名称。

例如：定义名称为 key123 的密钥链，并进入密钥链配置模式。

```
ruijie(config)#key chain key123
```

```
ruijie(config-keychain)#
```

② 定义密钥。在密钥链配置模式下使用 key 命令定义一个密钥，并进入密钥配置模式。

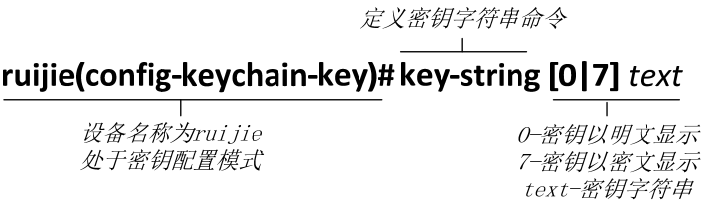
定义密钥的命令格式如下：

其中，*key-id* 为密钥链中验证密钥的 ID，取值范围为 0～2147483647。默认情况下，密钥链中无密钥。使用该命令的 no 形式删除指定密钥，如 ruijie(config-keychain)#no key。

例如：配置密钥链 ripkeys，并进入密钥链配置模式，配置密钥 1，并进入密钥 1 配置模式。

```
ruijie(config)#key chain ripkeys
ruijie(config-keychain)#key 1
ruijie(config-keychain-key)#
```

③ 定义密钥字符串。在密钥模式中使用 key-string 命令定义密钥字符串。该命令格式如下：



其中，“0”表示指定密钥以明文显示；“7”表示指定密钥以密文显示；*text* 表示指定的密钥验证字符串。默认情况下，密钥链无验证字符串。使用该命令的 no 形式删除密钥验证字符串，如 ruijie(config-keychain-key)#no key-string。

例如：配置密钥链 ripkeys，进入密钥链配置模式；配置密钥 1，进入密钥 1 配置模式；并定义密钥串为 abc。

```
ruijie(config)#key chain ripkeys
ruijie(config-keychain)#key 1
ruijie(config-keychain-key)#key-string abc
```

④ 关联密钥链。定义完密钥链后，需要将其应用到某接口中才能生效。在接口配置模式下执行如下命令指定使用的密钥链：

其中，*name-of-keychain* 为密钥链名称，指定 RIP 验证所使用的密钥链。默认情况下，未关联任何密钥链。使用该命令的 no 形式删除指定的密钥链，如 ruijie(config-if)#no ip rip authentication key-chain。

例如：配置密钥链名称为 ripkeys，并进入密钥链配置模式；配置密钥 1，进入密钥 1 配置模式，并定义密钥字符串为 abc。Serial 0/0 接口启用了 RIP 明文验证，关联的密钥链为 ripkeys。

```
ruijie(config)#key chain ripkeys
ruijie(config-keychain)#key 1
ruijie(config-keychain-key)#key-string abc
ruijie(config-keychain-key)#exit
ruijie(config-keychain)#exit
ruijie(config)#interface serial 0/0
ruijie(config-if)#ip rip authentication mode text
ruijie(config-if)#ip rip authentication key-chain ripkeys
```

（2）MD5 验证方式。

密文验证就是双方发起验证时验证密码以密文方式发送给对方。如果要使用 MD5 验证，则必须通过关联密钥链进行 MD5 验证。

密钥验证与明文验证过程一样，也是先定义密钥链，然后定义一个密钥 ID 及密钥字符串，最后在接口下关联该密钥链。关于定义密钥链、密钥 ID 及密钥字符串等命令，由于前面明文验证部分已详细介绍过，在此就不重复描述了。启动密钥验证唯一不同的地方就是定义验证方式时选择 MD5。定义 MD5 密文验证方式的命令格式如下：

```
ruijie(config-if)#ip rip authentication mode    md5
```

例如：配置密钥链 ripkeys，进入密钥链配置模式，配置密钥 1；进入密钥 1 配置模式，并定义密钥串为 abc。Serial 0/0 接口启用了 RIP 密文验证，关联的密钥链为 ripkeys。

```
ruijie(config)#key chain ripkeys
ruijie(config-keychain)#key 1
ruijie(config-keychain-key)#key-string abc
ruijie(config-keychain-key)#exit
ruijie(config-keychain)#exit
ruijie(config)#interface serial 0/0
ruijie(config-if)#ip rip authentication mode MD5
ruijie(config-if)#ip rip authentication key-chain ripkeys
```

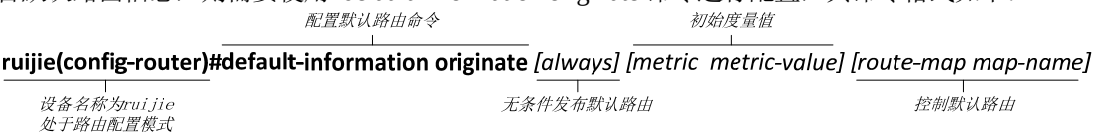
如果采用明文验证，但未配置明文验证字符串，或者未配置关联密钥链，或者关联了密钥链但密钥串实际未配置，此时并不会有验证行为发生。同样，如果采用 MD5 验证，但未配置关联密钥链，或者关联了密钥链但密钥串实际未配置，也不会有验证行为发生。

3）配置 RIP 默认路由

通过 RIP 动态路由协议能够学习到本 AS 内部的有限路由信息，使用 RIP 无法将网络（尤其 Internet 网络）中的全部路由学习到本路由器中。为了访问路由表之外的网络，需要配置默认路由，以便按照默认路由指向转发数据包。

在网络中配置默认路由有两种办法：一种是使用静态配置方法，在每台设备上分别配置，比较麻烦，冗余出错；另一种是在网关设备上使用默认路由命令发布默认路由。

如果路由设备的路由表中存在默认路由，RIP 默认情况就不会向外通告默认路由；如果欲向外通告默认路由信息，则需要使用 default-information originate 命令进行配置，其命令格式如下：



其中，*always* 表示如果选择 always 参数，RIP 路由进程不管是否存在默认路由，都会向邻居通

告一条默认路由，但该默认路由并不会在本地路由表中显示。

metric-value 为设置默认路由初始的度量值，metric-value 取值范围为 1～15，但优先级低于 route-map 的 set metric 规则；如果没有配置 metric 参数，则默认路由使用 RIP 配置的默认度量值。

map-name 为定义的 route-map 名称。如果要对 RIP 通告默认路由进行更多的控制，如使用 set metric 规则设置默认路由的度量值。

例如：通过配置命令向外发布一条默认路由，使得其他路由器获得一条默认路由。

```
ruijie(config-router)#default-information originate always
```

7.3　知识扩展

设置 RIP 定时器

RIP 分别有更新定时器（Update）、无效定时器（Invalid）、清除定时器（Flush）三个定时器。如图 7.3 所示，此三个定时器对 RIP 路由的更新与管理有着不同影响。

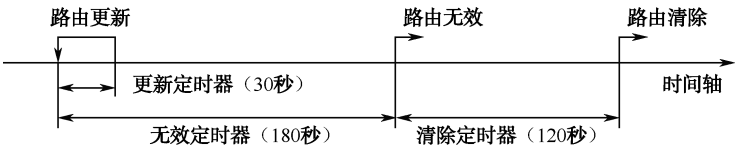


图 7.3　RIP 定时器功能

（1）更新定时器（Update）：定义路由更新时间，以秒计。Update 定义了设备发送路由更新报文的周期，每接收到更新报文，Invalid 和 Flush 时钟就复位。默认每隔 30 秒发送一次路由更新报文。

（2）无效定时器（Invalid）：定义路由无效时间，以秒计，从最近一次有效更新报文开始计时。Invalid 定义了路由表中路由因没有更新而变为无效的时间。路由无效时间至少应该为路由更新时间的 3 倍，如果在路由无效时间内没有接收到任何更新报文，相应的路由将变为无效，进入 Invalid 状态；Invalid 时间内接收到路由更新报文，时钟会复位。Invalid 默认时间为 180 秒。

（3）清除定时器（Flush）：定义路由清除时间，以秒计，从 RIP 路由进入 Invalid 状态开始计时。Flush 时间到期，处于 Invalid 状态路由将被清除出路由表。Flush 默认时间为 120 秒。

通过调整以上时钟，可能会加快路由协议的收敛时间及故障恢复时间。连接在同一网络上的设备，RIP 时钟值一定要一致。一般不建议对 RIP 时钟进行调整，除非有明确的需求。

修改 RIP 定时器的命令格式如下：

设置RIP定时器命令　　*无效时间*

ruijie(config-router)#timers basic *update invalid flush*

设备名称为ruijie
处于路由配置模式　　*更新时间*　　*清除时间*

其中，*update* 为路由更新时间，以秒计，默认值为 30 秒；*invalid* 为路由无效时间，以秒计，默认值为 180 秒；*flush* 为路由清除时间，以秒计，默认值为 120 秒。该命令的 no 形式恢复默认配置，如 ruijie(config-router)# no timers basic。

例如：设置路由器 RIP 更新报文每隔 10 秒发送一次；如果 30 秒内没有收到更新报文，相应的路由将变为无效路由，并进入 Invalid 状态；该路由进入 Invalid 状态后，超过 90 秒，将被清除。

```
ruijie(config)#router rip
```

```
ruijie(config-router)#timers basic 10 30 90
```

例如：将路由器 RIP 定时器恢复为默认值。

```
ruijie(config)#router rip
ruijie(config-router)#no timers basic
```

7.4 演示实例

1. 背景描述

网络中两台路由器通过以太网互连，运行 RIP 路由协议，采用明文验证方式，验证密钥为 key123，网络拓扑如图 7.4 所示。

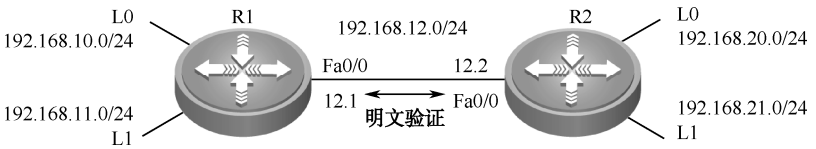


图 7.4 RIP 明文验证

2. 操作步骤

（1）路由器 R1 配置如下。

密钥链配置：

```
ruijie(config)#hostname R1
R1(config)#key chain ripkey
R1(config-keychain)#key 1
R1(config-keychain-key)#key-string key123
R1(config-keychain-key)#
```

配置以太网接口：

```
R1(config)#interface FastEthernet0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#ip rip authentication mode text
R1(config-if)#ip rip authentication key-chain ripkey
R1(config-if)#exit
R1(config)#interface loopback 0
R1(config-if)#ip address 192.168.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 1
R1(config-if)#ip address 192.168.11.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#
```

配置 RIP 路由协议：

```
R1(config)#router rip
R1(config-router)#network 192.168.10.0
```

```
R1(config-router)#network 192.168.11.0
R1(config-router)#network 192.168.12.0
R1(config-router)#version 2
R1(config-router)#end
R1#
```

（2）路由器 R2 的配置如下。

密钥链配置：

```
ruijie(config)#hostname R2
R2(config)#key chain ripkey
R2(config-keychain)#key 1
R2(config-keychain-key)#key-string key123
R2(config-keychain-key)#
```

配置以太网接口：

```
R2(config)#interface FastEthernet0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#ip rip authentication mode text
R2(config-if)#ip rip authentication key-chain ripkey
R2(config-if)#exit
R2(config)#interface loopback 0
R2(config-if)#ip address 192.168.20.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#interface loopback 1
R2(config-if)#ip address 192.168.21.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#
```

配置 RIP 路由协议：

```
R2(config)#router rip
R2(config-router)#network 192.168.20.0
R2(config-router)#network 192.168.21.0
R2(config-router)#network 192.168.12.0
R2(config-router)#version 2
R2(config-router)#end
R2#
```

（3）测试验证效果。

首先，通过 show ip route 命令查看路由器的路由表，应该能够看到另一台路由器的 loop0 及 loop1 网络的路由信息。

其次，可以将两个验证密码中的一个进行修改，使得两个验证密码不一致，如一个为 key123，另一个为 key456。稍过一段时间后，再次使用 show ip route 命令查看路由表，发现路由器应该学不到另一台路由器的 loop0 和 loop1 网络的路由信息了。

为什么学不到对方路由器的路由信息了呢？如果两台路由器的密钥链中有多个密钥（key），多个密钥中的密钥字符串相同或不同会出现什么情况呢？请大家认真思考。

最后，再次将两台路由器的验证密码恢复为 key123，稍过一段时间后，再次使用 show ip route 命令查看路由表，又能看到另一台路由器的 loop0 及 loop1 网络的路由信息了。

7.5 训练任务

1. 背景描述

网络中两台路由器通过以太网互连，运行 RIP 路由协议，采用 MD5 认证方式，验证密钥为 key123，网络拓扑如图 7.5 所示。

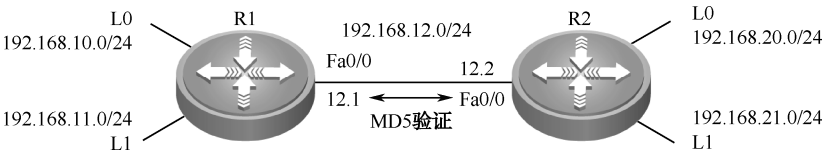


图 7.5 RIP 密文验证

2. 操作提示

将路由验证方式设置为 MD5 验证方式，路由器 R1 使用的命令如下：

```
R1(config-if)#ip rip authentication mode md5
```

所有其他配置命令都相同。

练习题

1. 选择题

（1）下面关于 RIP 版本 2 的描述错误的是（　　）。

A．RIPv2 为无类路由　　B．支持 VLSM

C．不支持自动汇总　　D．支持手动汇总

（2）默认情况下，RIP 路由协议每（　　）秒发送一次更新报文。

A．10　　B．20　　C．30　　D．40

（3）下面关于路由汇总功能的描述正确的是（　　）。

A．RIP 路由自动汇总的功能，提高了网络的伸缩性和有效性

B．RIP 路由自动汇总的功能，能够在路由表中看包含在汇总路由内的子路由

C．RIP 路由自动汇总的功能，可以避免造成路由黑洞

D．RIP 路由自动汇总的功能，使得查询路由表效率降低

（4）下面关于 RIP 验证的描述错误的是（　　）。

A．RIPv1 不支持验证

B．RIPv2 支持验证

C．RIPv2 支持两种验证方式：区域内验证和区域之间验证

D．RIPv2 支持两种验证方式：明文验证和 MD5 验证

（5）下面关于默认路由的描述错误的是（　　）。

A．默认路由是静态路由的一种特殊形式

B．默认路由是最后一个执行的路由

C．在网关设备上即使使用默认路由命令也不能发布默认路由

D．如果路由设备的路由表中存在默认路由，RIP 默认情况不会向外通告默认路由

2. 简答题

（1）描述路由汇总功能带来的好处。

（2）说明配置 RIP 密文验证的步骤。

（3）说明 RIP 定时器中更新定时器、无效定时器及清除定时器的含义及默认值。

项目 8　OSPF 路由协议与高级配置

在《网络设备配置与调试项目实训》中学习了 OSPF 路由协议的基本特性及配置，本章内容将在前面学习的基础上，进一步了解 OSPF 路由协议的高级特性及应用配置技能，包括 OSPF 路由汇总、OSPF 特殊区域、OSPF 虚拟链路、OSPF 验证等。

知识点、技能点

1. 掌握 OSPF 协议路由汇总功能及配置技能;
2. 掌握 OSPF 特殊区域特性及配置技能;
3. 掌握 OSPF 虚拟链路特性及配置技能;
4. 掌握 OSPF 验证功能及配置技能。

8.1　问题提出

某公司网络采用 OSPF 路由协议进行配置，但一方面由于子网数量多，且网络不够稳定，造成公司三层设备路由表波动大，设备负担重；另一方面，公司出于安全考虑，希望能够配置 OSPF 验证，实现只有合法设备才能够学习到路由信息。请您根据公司要求制定合适的网络实施方案，并完成相应配置工作。

8.2　相关知识

8.2.1　回顾 OSPF 协议的基本特性及应用配置

通过前面的学习我们已经了解 OSPF 的基本特性，下面我们再简单地回顾一下相关内容。

1. OSPF 协议特性

OSPF 路由协议是指开放最短路径优先（Open Shortest Path First）协议，是由 IETF（Internet Engineering Task Force）于 1988 年提出的一种链路状态路由选择协议，它服务于 IP 网络。OSPF 协议是内部网关协议 IGP 之一，它工作在一个自治系统 AS 内部用于交换路由选择信息。

OSPF 协议与前面介绍的 RIP 协议相比较，有如下不同之处。

（1）向谁发送路由选择信息?

OSPF 协议路由器通过输出接口向所有相邻路由器发送链路状态信息，每一台相邻路由器又将此信息转发给其他相邻路由器。这样，最终整个区域内的所有路由器都将得到这个信息。而 RIP 协议只向相邻路由器发送自己的路由表。

（2）发送什么样信息?

OSPF 协议路由器发送的内容是链路状态，包括链路度量值、邻居信息、UP 状态等。而 RIP 协议向自己邻居发送路由表。

（3）何时发送信息？

OSPF 协议只有当网络链路状态发送变化时，路由器才向所有邻接路由器用组播方式发送此变化的链路状态信息。而 RIP 协议不管网络拓扑是否变化，路由器之间都定期交换路由表信息。

2. OSPF 协议工作过程

（1）建立链路状态数据库 LSDB（Link-State Database）。如上所述，当网络拓扑发生变化后，检测到变化的路由器生成并发送链路状态通告 LSA（Link-State Advertisement），并通过组播地址发送给所有的邻居路由器。接收到 LSA 的每台路由器都复制一份 LSA，更新自己的链路状态库，然后再将 LSA 转发给其他的邻居。

（2）计算 SPF（Shortest Path First）树。通过使用 SPF 算法以本路由器为根计算到达所有其他目的网络的所有路径，形成一个 SPF 树。

（3）产生路由表。经过比较到达网络中所有目的网络的所有路径，选出最佳路径，产生路由表。

OSPF 协议根据网络的工作状况，经过建立 LSDB、计算 SPF 树、产生路由表等过程，能够快速地建立路由表，如图 8.1 所示。

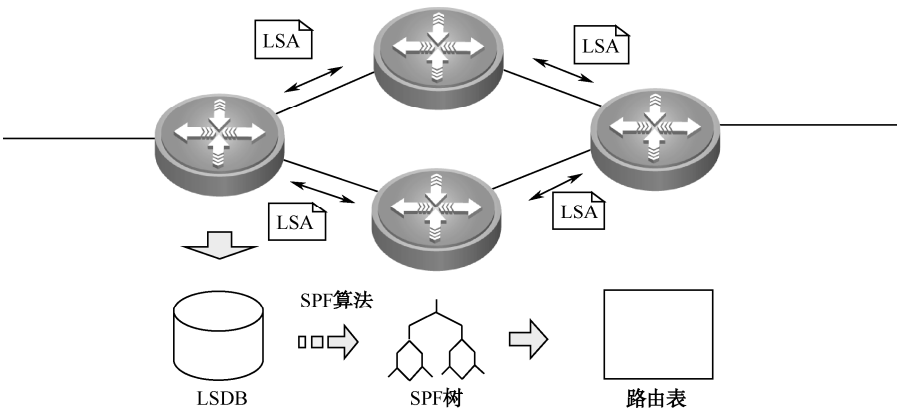


图 8.1　OSPF 协议工作过程

3. OSPF 网络分层结构

为了使 OSPF 能够用于规模很大的网络，OSPF 协议将自治系统 AS 划分为若干个更小的范围，每个范围也称为一个区域（Area）。OSPF 区域分成两种类型：骨干区域和非骨干区域。骨干区域连接其他非骨干区域，起到中枢传输作用；非骨干区域必须连接到骨干区域。每个区域都有一个 32 位的区域标志符，骨干区域的标志符为 0.0.0.0。

OSPF 网络划分区域后，网络中路由器根据位置不同其扮演的角色也不同。路由器可以分为 4 种类型：区域内路由器（Internal Router）、区域边界路由器 ABR（Area Border Router）、自治系统边界路由器 ASBR（Autonomous System Boundary Router）和骨干路由器（Backbone Router），如图 8.2 所示。

（1）区域内路由器。区域内路由器是指在一个 OSPF 区域内的路由器，这些路由器不与其他区域路由器相连，并且当网络链路发生变化时，只与区域内的其他路由器交换 LSA，也包括区域边界路由器，维护其所在区域的 LSDB，在本区域内实现收敛。如 R1 是区域内路由器。

（2）区域边界路由器 ABR。区域边界路由器 ABR 是指同时连接多个区域的路由器，ABR 维护多个区域的 LSDB。在 OSPF 网络中，所有的区域都必须与骨干区域（Area 0）相连， 因此 ABR 至少连接一个骨干区域和一个非骨干区域。如 R2、R3 是区域边界路由器。

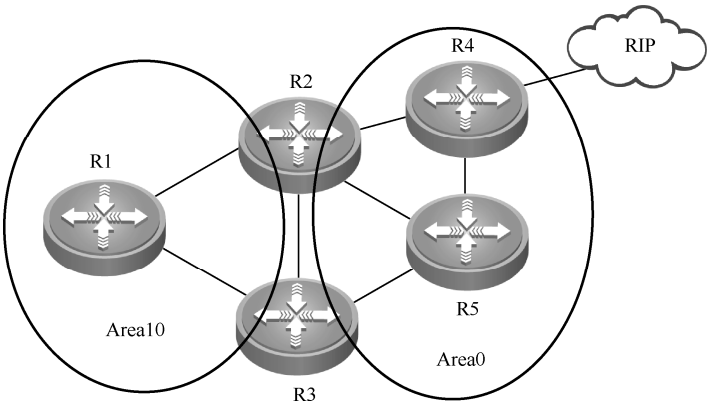


图 8.2　OSPF 分区示意图

（3）骨干路由器。骨干路由器是指在骨干区域（Area 0）的路由器，这些路由器只维护骨干区域的 LSDB。在 OSPF 网络中，所有非骨干区域之间的信息必须通过骨干区域进行转发。如 R5 是骨干路由器。

（4）自治系统边界路由器 ASBR。自治系统边界路由器 ASBR 是指与其他自治系统相连的路由器，通常这些路由器上使用多种路由协议（RIP、OSPF、BGP-4）。如 R4 除了是骨干路由器外还是自治系统边界路由器。

将自治系统划分为不同区域有如下优点。

（1）减少了整个网络的通信量。由于划分区域，每个区域内的路由器数量减少，相对来讲，链路状态的变化也减少了。属于同一个区域的路由器仅在该区域内互相发送链路状态通告 LSA，在区域内的路由信息流量也就减少了。

（2）减小了链路状态数据库 LSDB 的数据量。链路状态数据库用于描述该区域的拓扑结构，由于区域内链路减小了，因此链路状态数据库的数据量也减小了。同时，同一区域中的每台路由器都仅为该区域计算其 SPF 树。

（3）隐藏了区域内网络结构。每个路由器 LSA 只在区域内传播，对于外部区域路由器只能获得汇总信息，这样可以隐藏区域内网络的拓扑结构。

4. OSPF 协议报文格式

OSPF 不用 UDP 而是直接用 IP 数据报传送（其 IP 数据报首部的协议字段值为 89），如图 8.3 所示。

图 8.3　OSPF 分组用 IP 报文传送

OSPF 分组首部无论是哪种类型的分组都有着 24 字节的固定长度首部，但是分组的数据部分，5 种类型的 OSPF 分组各不相同。OSPF 分组首部格式如图 8.4 所示。数据部分稍后介绍。

其中，版本：用于 IPv4 的 OSPF 是版本 2。用于 IPv6 的 OSPF 是版本 3。

类型：定义 OSPF 数据报类型。OSPF 数据报类型共有 5 种：问候分组（Hello）、数据库描述分组（Database Description）、链路状态请求分组（Link-State Request）、链路状态更新分组（Link-State Update）、链路状态确认分组（Link-State Acknowledgement）。后续详细介绍。

图 8.4　OSPF 分组首部格式

分组长度：定义整个数据报长度，包括 OSPF 首部在内的分组长度，单位为字节。

路由器标志符（RID）：RID 是 OSPF 路由器的唯一标志符，称为路由器 ID。路由器 ID 的选举规则是，如果 loopback 接口不存在的话，就选举物理接口中 IP 地址等级最高的 IP 地址，否则就选举 loopback 接口。路由器 ID 对于建立邻居关系和协调 LSU 交换非常重要。在选举 DR、BDR 的过程中，如果 OSPF 优先级相同，则 RID 将用于决定谁赢得选举。如果该接口故障，该路由器就不可达。为了避免发生这种情况，最好定义一个回环接口 loopback 作为强制的 OSPF 路由器 ID。

区域标志符（Area ID）：为了能够通信，OSPF 路由器的接口必须属于一个相同的区域（Area），即共享子网及子网掩码信息。这些路由器拥有的链路状态必须相同。

检验和：用来检验分组中的差错。

验证类型：使用的验证模式。0 为没有验证；1 为简单口令验证；2 为加密检验和验证（MD5）。

验证：报文验证的必要信息。当验证类型为 0 时，不检验该字段；当验证类型为 1 时，该字段包含一个最长为 64 位的口令；当验证类型为 2 时，这个字段包含一个 key ID、验证数据长度和一个加密序列号。

如果验证类型为 2，验证字段（64bit）内容如图 8.5 所示。

图 8.5　OSPF 分组验证类型=2 的验证字段格式

5. OSPF 分组类型

在 OSPF 所有数据报中，首部格式都是相同的，主要区别就是数据部分因 OSPF 分组类型不同而不同。OSPF 分组类型共有 5 种：问候分组（Hello）、数据库描述分组（Database Description）、链路状态请求分组（Link-State Request）、链路状态更新分组（Link-State Update）、链路状态确认分组（Link-State Acknowledgement）。下面分别介绍每种类型数据报的格式。

（1）问候分组（Hello）。Hello 分组用来建立和保持 OSPF 邻接关系。由于 OSPF 协议根据路由器之间的链路状态进行路由选择，Hello 采用多播地址 224.0.0.5，确保邻居之间的双向通信来建立和维护邻接关系。

当路由器在从邻居那里收到的 Hello 报文中“看到”自己后，便进入了双向通信状态，建立了邻接关系，Hello 分组格式如图 8.6 所示。

图 8.6　Hello 分组格式

在 Hello 分组数据部分中，各字段功能描述如下。

① Hello 间隔时间：表示发送两个 Hello 分组之间相隔的时间秒数。这个间隔对于两个正尝试形成一个邻接体关系的路由器来说必须是相同的。Hello 间隔在广播介质和点到点介质中都是 10 秒，而在其他介质中是 30 秒。

② 选项：表示路由器支持的可选特性。

③ 路由器优先级：默认时，这个值被设置为 1。这个字段在选举 DR 和 BDR 的时候扮演着重要角色。高的优先级增加了这台路由器变成 DR 的机会。优先级为 0 表示这台路由器将不参与 DR 的选举。

④ 失效时间：表示在一个邻居被宣布死亡之前以秒为单位的时间数目。默认时，死亡间隔是 Hello 间隔的 4 倍。

⑤ DR：列出指定路由器 DR 的 IP 地址。如果没有 DR，则这个字段值为 0.0.0.0。DR 是通过 Hello 协议选举出来的，有着最高优先级的路由器变成 DR。如果优先级相等，有着最大路由器 ID 的路由器成为 DR。DR 的目的是在多点接入介质（Multi Access Media）中减少泛洪的数量。DR 使用多播来减少泛洪的数量。所有的路由器将它们的链路状态数据库向 DR 泛洪，同时 DR 又将这些信息反过来泛洪给这个网段中其他的路由器。在点到点或者点到多点网段中没有 DR 或者 BDR 存在。

⑥ BDR：列出备份指定路由器 BDR 的 IP 地址。如果 BDR 不存在，这个字段值为 0.0.0.0。BDR 也是通过 Hello 协议选举出来的。BDR 的目的是作为 DR 的备份，在 DR 死亡的时候作一个平滑的转换。BDR 在泛洪中保持被动。

⑦ 邻居列表：邻居列表字段中包含已建立双向通信关系的邻居路由器 ID。路由器在邻居发送的 Hello 分组中的邻居列表字段中“看到”自己后，便表明双向通信关系已经建立。

（2）数据库描述分组（Database Description）。当两台 OSPF 路由器初始化连接时，要交换数据库描述 DBD 分组。这个 DBD 分组用于描述链路状态数据库内容。由于链路状态数据库的内容可能很多，所以可能需要几个数据库描述分组来描述数据库，这些描述分组有着专用的数据库描述分组序列字段。

数据库描述分组的格式如图 8.7 所示。

图 8.7　数据库描述分组格式

① Interface MTU：表示接口最大传输数据单元，以字节为单位。用来检查两端 OSPF 路由器接口的 MTU 是否匹配。在 Virtual-link 中的 Interface MTU 字段为 0。

② I 选项：当设置为 1 时，表示这是 DBD 交换中的第一个分组。在交换 DBD 分组时，需要协调主从，比较 Router-ID，Router-ID 值大的路由器为主路由器 Master。主路由器发送序列号，从路由器进行确认。

③ M 选项：当设置为 1 时，表示后面将有更多的分组过来。

④ M/S 选项：用于主从设备。当这个比特设置为 1 时，表示路由器在 DBD 交换过程中是主设备，如果这个比特被设置为 0，则表示路由器是从设备。

⑤ DD Sequence Number：这个字段是数据库描述分组序列号，包含一个由主设备设置的唯一的值，这个序列号在数据库交换过程中使用，只有主设备才能增加序列号。

⑥ An LSA Header：这个字段由一系列链路状态通告 LSA 报文头组成。交换 DBD 的目的是了解对方都拥有哪些链路状态通告 LSA 信息，所有 DBD 分组中的 LSA 并不是具体的 LSA 报文，而是每个 LSA 报文首部信息。LSA 首部信息用于唯一地标志一个 LSA。当发现需要哪个 LSA 时，才在后续开始请求交换该 LSA 完整信息。

（3）链路状态请求分组（Link-State Request）。链路状态请求分组用于请求相邻路由器的链路状态数据库中的信息。当路由器收到一个数据库描述 DBD 分组时，可以发现自己路由器中所缺少的信息。这样，路由器会发送一个或几个链路状态请求分组给邻居路由器以得到更多的链路状态信息。链路状态请求分组的格式如图 8.8 所示。

① LS 类型（LS type）：表示请求的 LSA 类型号（1～5）。

② 链路状态（Link State）ID：用于确定 LSA 描述的 OSPF 域部分。

③ 通告路由器（Advertisting Router）：初始建立 LSA 的路由器的路由器 ID。

（4）链路状态更新分组（Link-State Update）。链路状态更新分组用于把 LSA 发送给它的邻居，这些更新分组用于对 LSA 请求的应答。一个 LSU 中可以包括多个 LSA 条目，通常有 5 种不同类型的 LSA 分组，这些分组类型用从 1 到 5 的类型号来标志。链路状态更新分组的格式如图 8.9 所示。

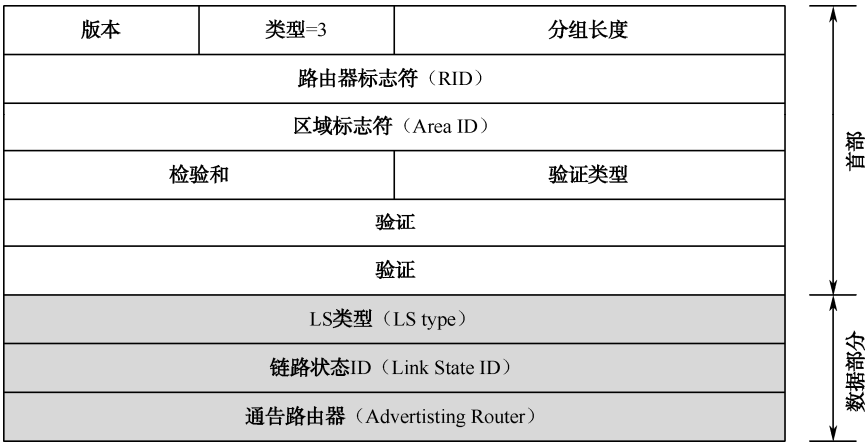


图 8.8　链路状态请求分组格式

图 8.9　链路状态更新分组格式

（5）链路状态确认分组（Link-State Acknowledgement）。链路状态确认分组用于当路由器收到对方 LSA 后，向对方发送的响应报文，实现 LSA 分组的可靠传输。

链路状态确认分组包含 LSA 头中的信息，如 Link State ID、通告路由器和 LS 顺序号等。链路状态确认分组与 LSA 间无须一对一的应用关系。一个链路状态确认分组可以包含对多个 LSA 的应答。链路状态确认分组的格式如图 8.10 所示。LSA 头格式后续介绍。

6. 链路状态通告 LSA（Link State Advertisement）格式

（1）LSA 报文头格式。在 OSPF 中有 5 种类型 LSA，所有的 LSA 报文的首部（报文头）都相同。一个 LSA 报文首部唯一地标志了一个 LSA 报文。LSA 报文首部的格式如图 8.11 所示。

① LS age：LSA 的年龄。表示自从 OSPF 产生时已消逝的秒数。

② Options：选项。由一系列标志组成，这些标志标识了 OSPF 网络能提供的各种可选的服务。

③ LS type：LS 类型，指出 5 种 LSA 类型中的一种。每种类型的格式各不相同。

④ Link State ID：链路状态 ID。表示 LSA 描述的特定网络环境。链路状态 ID 与 LS 类型密切相关，不同 LS 类型，表示链路状态 ID 方式也不同。如 LS 类型为路由器 LSA 时，链路状态 ID 用产生该 LSA 的路由器 ID 表示。

图 8.10　链路状态确认分组格式

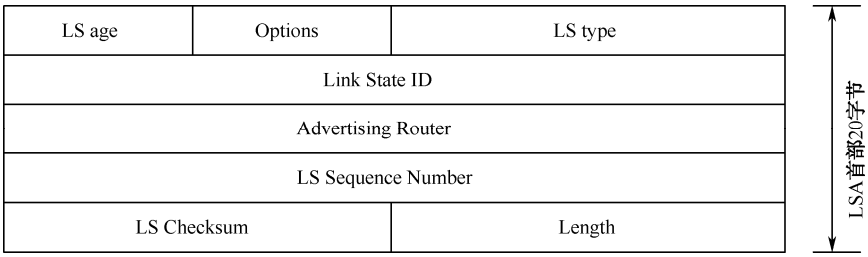


图 8.11　LSA 报文头格式

⑤ Advertising Router：通告路由器，表示产生了该 LSA 的路由器的 Router ID。

⑥ LS Sequence Number：LSA 序列号。OSPF 路由器为每个 LSA 编制一个序列号，并且每增加一个 LSA 就递增该序列号。通过序列号可以判断一个 LSA 是否是新的 LSA。

⑦ LS Checksum：LS 检验和。用于检测 LSA 在传输过程中是否受到破坏。

⑧ Length：LS 长度。表示 LSA 长度，以字节为单位。

（2）LSA 报文分类。下面分别说明类型 1～5 的 LSA 报文。

① 路由器 LSA（Router LSA，类型 1）：由 OSPF 路由器产生的向所连接区域发送的路由器 LSA，它描述了路由器到区域的链路状态，如接口的花费值、接口的地址等信息。路由器必须为它属于的每个区域产生一个路由器 LSA，所以区域边界路由器将产生多个路由器 LSA。路由器 LSA 只在一个区域内传播，不会穿越 ABR，如图 8.12 所示。

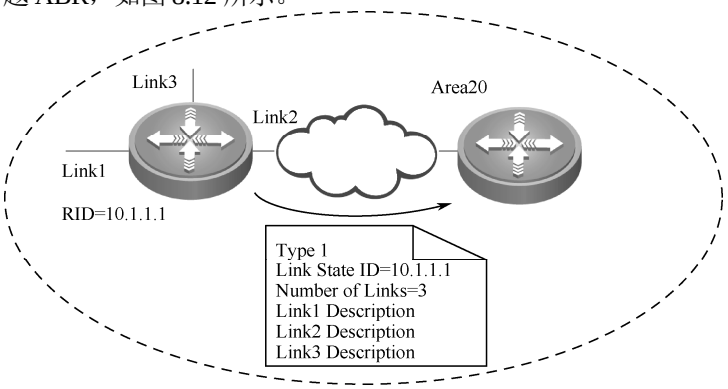


图 8.12　路由器 LSA

② 网络 LSA（Network LSA，类型 2）：指定路由器（DR）产生的将连接到某个网段的所有路由器的链路状态和花费值向多接口网络及所有连接在其上的路由器发送的网络 LSA。网络 LSA 减少网络中的路由更新信息流量。网络 LSA 只在一个区域内传播，不会穿越 ABR，如图 8.13 所示。

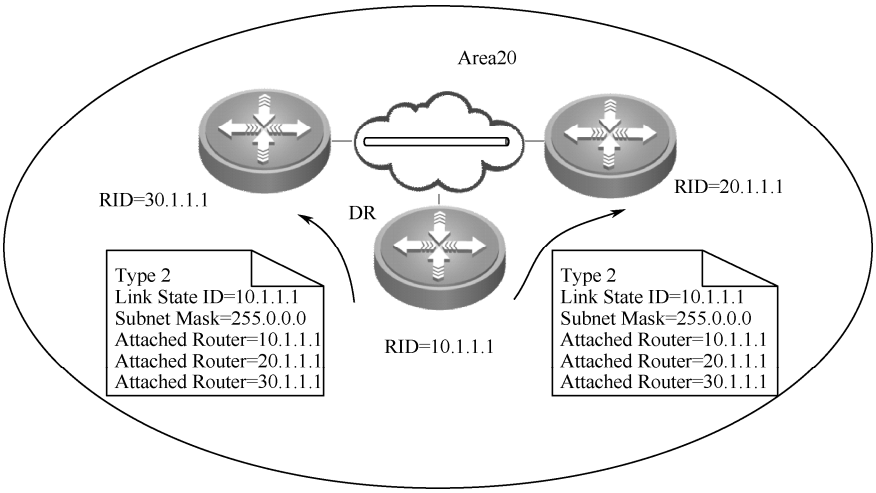


图 8.13　网络 LSA

③ 网络汇总 LSA（Network Summary LSA，类型 3）：由区域边界路由器 ABR 产生的将一个区域的路由信息或汇总后的路由信息向另一个区域发送的网络汇总 LSA。网络汇总 LSA 实现了一个自主系统内的不同区域（Area）间路由共享。网络汇总 LSA 在整个 OSPF 域内泛洪，如图 8.14 所示。

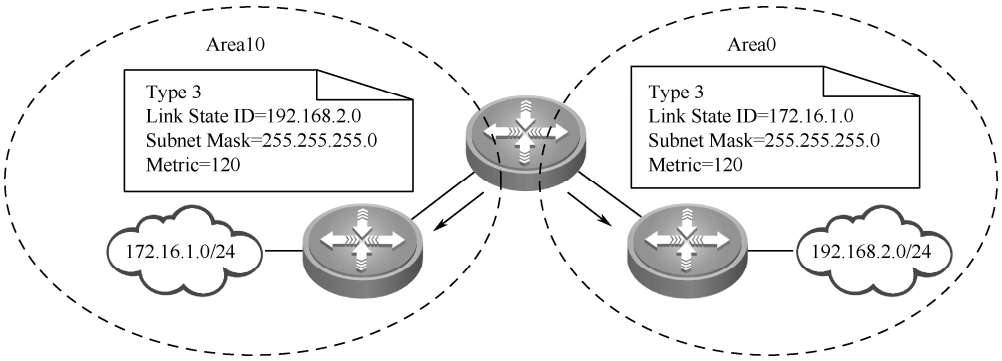


图 8.14　网络汇总 LSA

④ ASBR 汇总 LSA（ASBR Summary LSA，类型 4）：由区域边界路由器 ABR 产生的在 OSPF 域内发送 ASBR 位置的 ASBR 汇总 LSA。ASBR 汇总 LSA 实现 OSPF 域内路由器访问 OSPF 域外网络时的出口路由信息。ASBR 汇总 LSA 在整个 OSPF 域内泛洪，如图 8.15 所示。

⑤ 自治系统外部 LSA（AS-External LSA，类型 5）：由 ASBR 产生的用于描述 OSPF 网络之外的目的地自主系统外部 LSA。自治系统外部 LSA 实现了把 OSPF 外部网络路由信息在整个 OSPF 路由域内传播。自治系统外部 LSA 在整个 OSPF 域内泛洪（除了 Stub 区域），如图 8.16 所示。

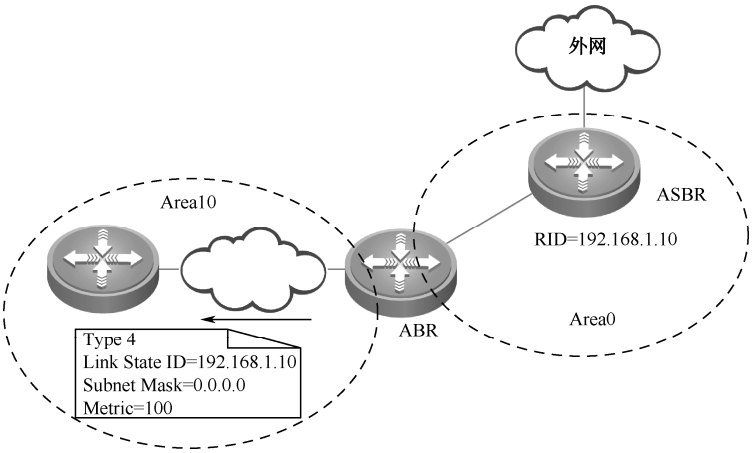


图 8.15　ASBR 汇总 LSA

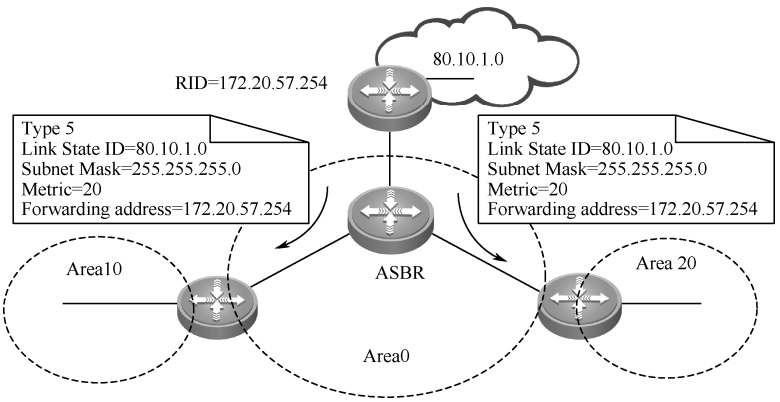


图 8.16　自治系统外部 LSA

7. OSPF 的建立邻居状态与数据库同步过程

（1）建立双向通信。运行 OSPF 协议的路由器初始化时，首先使用 Hello 分组完成如图 8.17 所示交换过程。

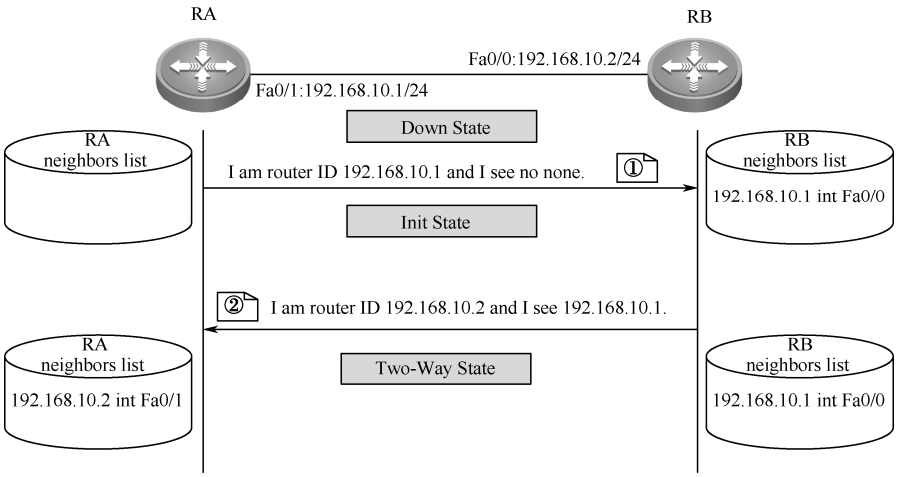


图 8.17　建立双向通信过程

交换过程说明如下。

① 初始路由器 RA 处于 Down 状态下，它首先从其 OSPF 接口向外发送 Hello 报文，发送 Hello 报文使用的多播地址为 224.0.0.5。

② 路由器 RB 收到 RA 发送的 Hello 报文，把路由器 RA 加进自己的邻居列表（Neighbors list）中，并进入初始化状态（Init State）。以单播的方式发送 Hello 报文以对路由器 RA 应答。

③ 路由器 RA 收到 RB 发送的 Hello 报文后，将路由器 RB 加进自己的邻居列表（Neighbors list），并进入双向通信状态（Two-Way State）。

④ 如果链路是广播型网络，如以太网，则接下来进行 DR 和 BDR 选举。这一步发生在交换信息之前。

⑤ 周期发送 Hello 报文保持信息交换，路由器每隔 10 秒交换一次 Hello 报文。

（2）选举 DR 和 BDR。在广播型的 OSPF 网络中，为了减少路由更新流量，需要选举指定路由器 DR 和备份指定路由器 BDR。网络中的每台路由器都必须与 DR 和 BDR 建立邻居关系。网络中的路由器只将链路信息发送至 DR 和 BDR，而不是发送给所有其他路由器。DR 收到路由器发送的链路状态信息后，将其转发给网络中的其他路由器，如图 8.18 所示。

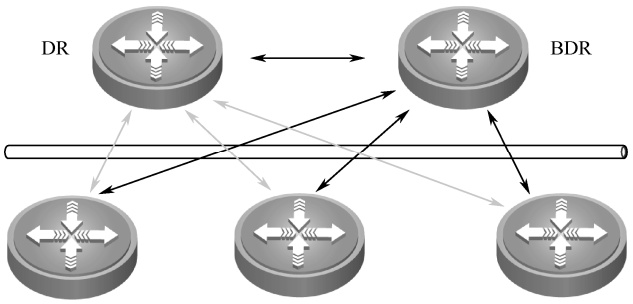


图 8.18　DR 和 BDR

非 DR 和 BDR 路由器将链路状态信息发送至 DR 使用多播地址 224.0.0.6，而 DR 使用多播地址 224.0.0.5 再将链路状态信息转发给非 DR 和 BDR 路由器。

当选举 DR 和 BDR 时，需要比较 Hello 报文中优先级（Prioty），优先级高的为 DR，次高的为 BDR，默认优先级为 1。在优先级相同情况下，比较 RID，RID 等级最高的为 DR，次高的为 BDR。当优先级设置为 0 时，OSPF 路由器将不能成为 DR 或 BDR，只能成为 DROther。

（3）发现网络路由及添加链路项目。当选举完 DR 和 BDR 后，进入 Exstart 状态，接下来可以对链路状态信息进行发现并创建自己的 LSDB。

发现与添加路由过程如下。

① 在 Exstart 状态下，路由器和 DR/BDR 形成主从关系，以 RID 等级高的为主，RID 等级低的为从。

② 主从交换 DBD 报文，路由器进入 Exchange 状态，如图 8.19 所示。

③ DBD 报文包含了出现在 LSDB 中的 LSA 条目头部信息，每个 LSA 条目头部信息包括链路状态类型、通告路由器的地址、链路耗费和序列号等。

④ 路由器收到 DBD 报文后，将使用 LSAck 报文响应。同时将比较收到的 DBD 报文中的条目和自己的 DBD 条目。

⑤ 比较结果发现收到的 DBD 报文中有更新的条目，路由器就发送 LSR 报文给其他路由器，进入 Loading 状态。对方收到 LSR 报文后，对方路由器以发送 LSU 报文作为响应。LSU 报文包含了 LSR

中所需要的完整信息。收到 LSU 后，以再次发送 LSAck 做出确认。

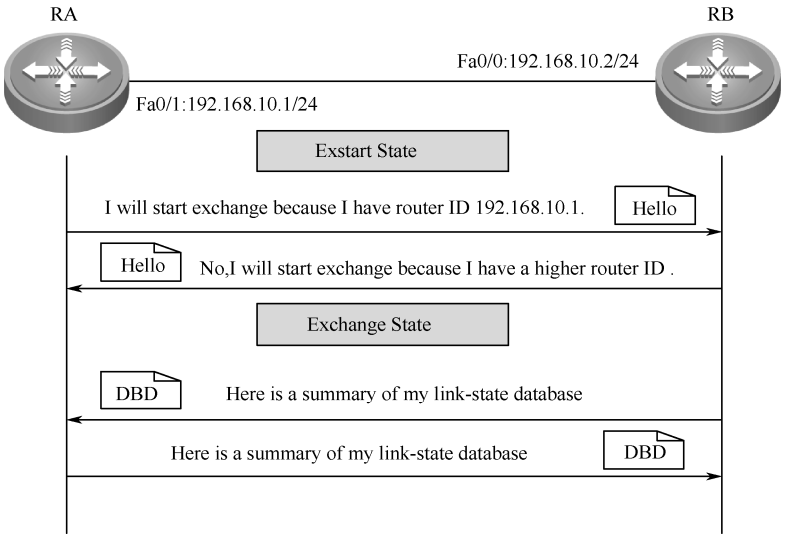


图 8.19　发现网络路由

⑥ 路由器添加新的条目到 LSDB 中，进入 Full 状态，至此，区域内的所有路由器的 LSDB 都相同，如图 8.20 所示。

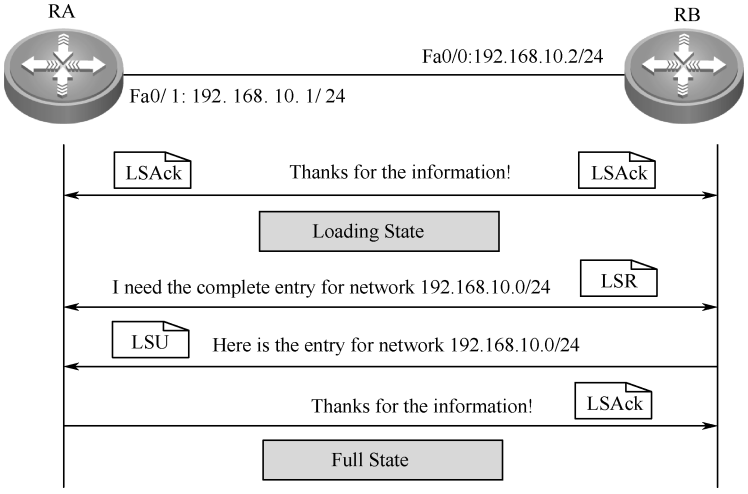
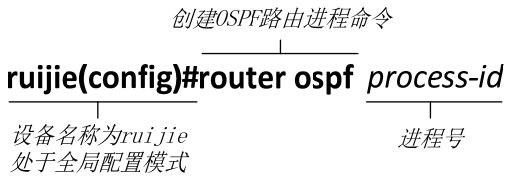


图 8.20　添加网络路由

网络中路由器的链路状态数据库 LSDB 达到同步后，OSPF 路由协议通过最短路径算法 SPF，产生路由表。OSPF 路由协议的链路状态数据库能够较快地更新，使各台路由器的路由表也得到及时更新，实现 OSPF 路由协议快速收敛。

8. OSPF 路由协议的基本配置

（1）创建 OSPF 路由进程。创建 OSPF 路由进程，并定义与该 OSPF 路由进程关联的 IP 地址范围，以及该范围 IP 地址所属的 OSPF 区域。OSPF 路由进程只在属于该 IP 地址范围的接口发送、 接收 OSPF 报文，并且对外通告该接口的链路状态。创建 OSPF 路由进程的命令格式如下：



其中，*process-id* 为 OSPF 进程号，本地路由器有效，可以是任意整数。默认情况下，不存在 OSPF 路由进程。使用该命令的 no 形式删除已定义的 OSPF 路由进程，如 ruijie(config)# no router ospf 1，删除进程号为 1 的 OSPF 路由进程。

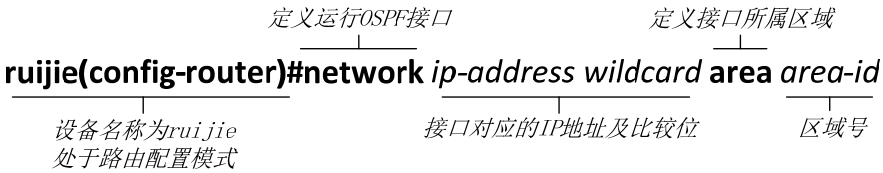
例如：创建 OSPF 路由进程 10。

```
ruijie(config)#router ospf 10
ruijie(config-router)#
```

例如：删除创建的 OSPF 路由进程 10

```
ruijie(config)#no router ospf 10
```

（2）定义 OSPF 接口及所属区域。定义运行 OSPF 的接口及这些接口所属区域，使用的命令格式如下：



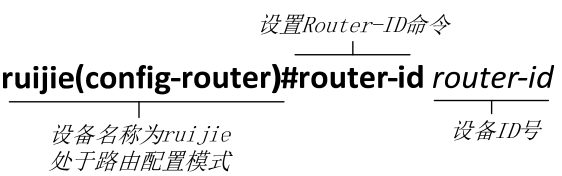
其中，*ip-address* 为接口对应的 IP 地址；*wildcard* 为 IP 地址比较比特位，“0”表示精确匹配，“1”表不做比较；*area-id* 为 OSPF 区域标志。一个 OSPF 区域总是关联一个地址范围，为了便于管理，也可以用一个子网作为 OSPF 区域标志。

默认情况下，没有配置 OSPF 区域。使用该命令的 no 形式删除接口的 OSPF 区域定义，如 ruijie(config-router)# no network 192.168.1.0 0.0.0.255 area 10，删除在区域 10 中已定义的地址为 192.168.1.0 的接口。

例如：在某网络中，定义了 3 个区域：0，1，172.16.16.0。将 IP 地址落在 192.168.12.0/24 范围内的接口定义到区域 1，将 IP 地址落在 172.16.16.0/20 范围内的接口定义到区域 172.16.16.0，将其余接口定义到区域 0。

```
ruijie(config)#router ospf 20
ruijie(config-router)#network 172.16.16.0    0.0.15.255 area 172.16.16.0
ruijie(config-router)#network 192.168.12.0    0.0.0.255 area 1
ruijie(config-router)#network    0.0.0.0    255.255.255.255 area 0
```

（3）设置 Router-ID。Router-ID 是网络中 OSPF 设备的唯一标志，OSPF 路由进程启动时将选择 Router-ID。可以配置任何一个 IP 地址作为该 Router-ID，但是每台设备的 ID 标志必须唯一。可以通过命令修改 Router-ID，但是必须在启动路由进程之前设置 Router-ID。设置路由器 ID 的命令格式如下：



其中，*router-id* 为要设置的路由设备的 ID，以 IP 地址形式表示。默认情况下由 OSPF 路由进程选

举接口最大的 IP 地址作为路由设备标志。使用该命令的 no 选项可以删除所设置的 RouterID，恢复使用默认的 Router ID，如 ruijie(config-router)#no router-id。

例如：将路由器的 Router-ID 设置为为 0.0.0.36。

```
ruijie(config)#router ospf 20
ruijie(config-router)#router-id 0.0.0.36
```

（4）配置 OSPF 单区域实例。某网络拓扑如图 8.21 所示。路由器 R1 和 R2 分别启用 OSPF 路由协议，并且设置区域 0。

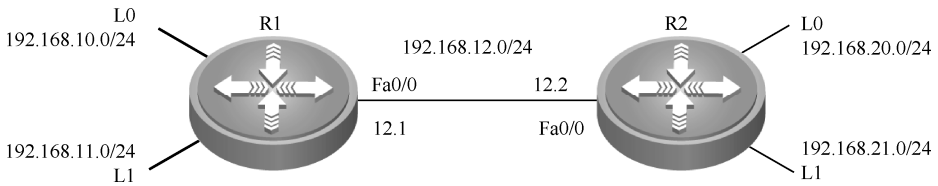


图 8.21　单区域 OSPF 网络

路由器 R1 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R1
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 0
R1(config-if)#ip address 192.168.10.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface loopback 1
R1(config-if)#ip address 192.168.11.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#router ospf 1
R1(config-router)#network 192.168.10.0 0.0.0.255 area 0
R1(config-router)#network 192.168.11.0 0.0.0.255 area 0
R1(config-router)#network 192.168.12.0 0.0.0.255 area 0
R1(config-router)#end
R1#
```

路由器 R2 配置代码如下：

```
ruijie#configure
ruijie(config)#hostname R2
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#ip rip send version 1 2
R2(config-if)#ip rip receive version 1 2
```

```
R2(config-if)#exit
R2(config)#interface loopback 0
R2(config-if)#ip address 192.168.20.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#interface loopback 1
R2(config-if)#ip address 192.168.21.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
R2(config)#router OSPf 1
R2(config-router)#network 192.168.12.0 0.0.0.255 area 0
R2(config-router)#network 192.168.20.0 0.0.0.255 area 0
R2(config-router)#network 192.168.21.0 0.0.0.255 area 0
R2(config-router)#end
R2#
```

使用 show ip route 命令查看路由器 R1、R2 的路由表中是否学习到相应的路由信息。

8.2.2 OSPF 协议特性及高级配置

1. OSPF 路由汇总

路由汇总是将多条路由汇总成一条路由。路由汇总的目的主要是减小路由表的规模、减少路由信息的通告数量，并且增强网络的稳定性。如果不进行路由汇总，每条详细的路由都将传播到 OSPF 骨干网络中，导致不必要的网络数据流量和系统开销。

OSPF 路由汇总分为区域间路由汇总和外部路由汇总两种。

（1）OSPF 区域间路由汇总。区域间路由汇总是在 ABR 上进行的，用来将一个区域的多个路由合并或汇聚成一条路由，然后通告到其他区域去。路由信息的合并行为只发生在区域边界，在区域内部路由设备上看到的都是具体的路由信息，但在区域外部只能看到一条汇总路由。可以定义多个区域路由汇总命令，这样整个 OSPF 路由域的路由将得到简化，特别是网络规模较大时，会提高网络转发性能。

配置 OSPF 区域间路由汇总的命令格式如下：

定义区域　　*定义是否要公布该汇聚范围*

ruijie(config-router)#area *area-id* **range** *ip-address net-mask* **[advertise | not-advertise] [cost** *cost***]**

设备名称为ruijie 处于路由配置模式　　*定义汇总路由范围*　　*设置汇总路由度量值*

其中，*area-id* 为指定被汇总路由的 OSPF 区域号；*ip-address* 为定义汇总后的网络地址；*net-mask* 为定义汇总后的子网掩码；*advertise* 为公布该汇总范围，默认情况为公布，设置该选项将为其产生一个类型 3 的 LSA；*not-advertise* 为不公布该汇总范围，设置该选项将不会为其产生一个类型 3 的 LSA；*cost* 为设置汇总路由的度量值。

默认情况下，没有配置区域间路由汇总。该命令的 no 形式删除已配置的路由汇总，no 前缀与 cost 参数组合可以恢复路由汇总默认的度量值，但不会删除路由汇总。如 ruijie(config-router)# no area 10 range 172.16.0.0 255.255.0.0 删除已定义的汇总 172.16.0.0 的功能。

例如：将区域 1 的路由汇总成一条路由 172.16.16.0/20

```
ruijie(config)#router ospf 1
```

```
ruijie(config-router)#network 172.16.0.0 0.0.15.255 area 0
ruijie(config-router)#network 172.16.17.0 0.0.15.255 area 1
ruijie(config-router)#area 0 range 172.16.16.0 255.255.240.0
```

（2）OSPF 外部路由汇总。外部路由汇总是在 ASBR 上进行的，用来将重分发到 OSPF 路由域的多条外部路由汇总成一条路由。当路由从其他路由进程重新分布，并注入 OSPF 路由进程时，每条路由均以一个外部链路状态的方式通告给 OSPF 路由设备。如果注入的路由是一个连续的地址空间，自治域边界路由器可以只通告一个汇总路由，从而大大减小路由表的规模。

要配置外部路由汇总，在路由进程配置模式中执行以下命令：

外部路由汇总命令　　*不公布该汇总路由*

ruijie(config-router)# summary-address *ip-address net-mask* **[not-advertise]**

设备名称为ruijie
处于路由配置模式　　*汇总地址及子网掩码*

其中，*ip-address* 为汇总后的 IP 地址；*net-mask* 为汇总后的网络掩码；not-advertise 为不公告该汇总路由，若未配置则为公告。

默认情况下，没有配置路由汇总。使用该命令的 no 形式删除路由汇总的定义。

例如：产生一条外部汇总路由 100.100.0.0/16。

```
ruijie(config)#router ospf 1
ruijie(config-router)#summary-address 100.100.0.0 255.255.0.0
ruijie(config-router)#redistribute static subnets
ruijie(config-router)#network 200.2.2.0 0.0.0.255 area 1
ruijie(config-router)#network 172.16.24.0 0.0.0.255 area 0
ruijie(config-router)#area 1 nssa
```

例如：某网络中，OSPF 区域 1 包含路由器 RA 和 RB，OSPF 区域 0 包含路由器 RB 和 RC，RB 作为 ABR 路由器；RC、RD 启用了 RIP 路由协议，RC 作为 ASBR 路由器。路由器 RA 中有 172.16.10.0、172.16.11.0、172.16.12.0 三个子网，可以在路由器 RB 上通过区域间路由汇总将区域 1 中的三条路由条目汇总成一条路由通告到区域 0 中；路由器 RD 也有 10.10.10.0、10.10.11.0、10.10.12.0 三个子网，可以在路由器 RC 上通过外部路由汇总将 RIP 中的三条路由汇总成一条路由通告到 OSPF 中，如图 8.22 所示。

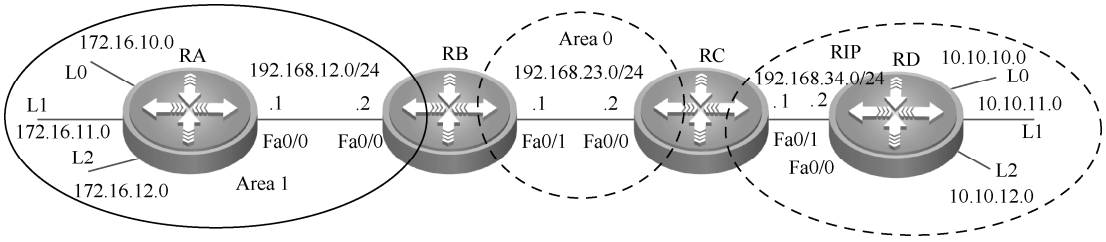


图 8.22　外部路由汇总

路由器 RA 及 RB 配置与前面配置相同，在此略。

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
```

```
RC(config-if)#ip address 192.168.23.2 255.255.255.0
RC(config-if)#exit
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 192.168.34.1 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 192.168.23.0 0.0.0.255 area 0
RC(config-router)#redistribute connected
RC(config-router)#redistribute rip metric 50 subnets
RC(config-router)#summary-address 10.10.10.0 255.255.248.0
RC(config-router)#exit
RC(config)#router rip
RC(config-router)#version 2
RC(config-router)#network 192.168.34.0
RC(config-router)#no auto-summary
RC(config-router)#redistribute connected
RC(config-router)#redistribute ospf metric 1
RC(config-router)#end
RC#
```

路由器 RD 配置代码如下：

```
RD#configure terminal
RD(config)#interface loop0
RD(config-if)#ip address 10.10.10.1 255.255.248.0
RD(config-if)#exit
RD(config)#interface loop1
RD(config-if)#ip address 10.10.11.1 255.255.248.0
RD(config-if)#exit
RD(config)#interface loop2
RD(config-if)#ip address 10.10.12.1 255.255.248.0
RD(config-if)#exit
RD(config)#interface fastethernet 0/0
RD(config-if)#ip address 192.168.34.2 255.255.255.0
RD(config-if)#exit
RD(config)#router rip
RD(config)#version 2
RD(config-router)#network 10.0.0.0
RD(config-router)#network 192.168.34.0
RD(config-router)#end
RD#
```

2. OSPF 验证

OSPF 路由协议同 RIP 一样，也可以通过验证功能验证邻居路由器的身份，避免路由信息外泄。OSPF 进行身份验证步骤也分为定义接口的验证方式及设置验证密码等步骤。

第 1 步：定义 OSPF 验证方式。

OSPF 身份验证方式既可以在区域内设置也可以在相应的接口上设置。如果在区域内设置了某种身份验证方式，那么在该区域内的所有路由器接口都需要进行验证，验证通过的接口才能传递 OSPF 链路状态信息。如果只需要在某接口进行身份验证，则可以采用接口身份验证方式，只在此接口进行身份验证。下面分别介绍这两种身份验证方式的设置命令格式。

（1）定义接口的 OSPF 验证方式的命令格式如下：

其中，*message-digest* 表示在该接口上进行 MD5 加密验证；*null* 表示在该接口上取消验证。

默认情况下，接口上没有设置验证方式，此时接口上采用的是所在的区域的验证类型。使用 no 选项可以将验证方式恢复为默认验证方式，如 ruijie(config-if)# no ip ospf authentication。

例如：某网络中，设置路由器 FastEthernet 0/0 接口的 OSPF 验证方式为 MD5 验证。

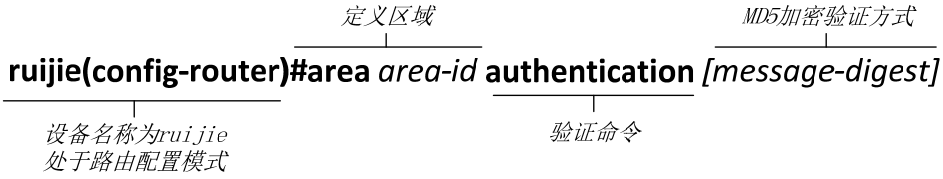
```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip address 172.16.1.1 255.255.255.0
ruijie(config-if)#ip ospf authentication message-digest
```

例如：取消已设置路由器 FastEthernet 0/0 接口的 OSPF 验证方式 MD5 验证，将其恢复为默认值。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip address 172.16.1.1 255.255.255.0
ruijie(config-if)#no ip ospf authentication
```

提示：*如果命令后面不跟任何选项表示进行明文验证。当接口上的验证方式和接口所在的区域的验证方式都设置时，优先采用接口上的验证方式。*

（2）设置 OSPF 区域验证方式的命令格式如下：



其中，*area-id* 指定要启用 OSPF 验证的区域号；*message-digest* 表示采用 MD5 验证方式；没有使用 message-digest 选项时，表示采用明文验证方式。

默认情况下，没有采用验证。使用该命令的 no 形式关闭 OSPF 区域验证，如 ruijie(config-rotuer)# no area area-id authentication。

例如：在某网络中，设置 OSPF 路由进程的区域 0（骨干区域）采用 MD5 验证方式，验证密码为 key123。

```
ruijie(config)#interface FastEthernet 0/0
ruijie(config-if)#ip address 192.168.12.1 255.255.255.0
ruijie(config-if)#ip ospf message-digest-key 1 md5 key123
ruijie(config)#router ospf 1
ruijie(config-router)#network 192.168.12.0 0.0.0.255 area 0
```

```
ruijie(config-router)#area 0 authentication message-digest
```

例如：在某网络中，设置 OSPF 路由进程的区域 0（骨干区域）采用明文验证方式，验证密码为 key123。

```
ruijie(config)#interface FastEthernet 0/0
ruijie(config-if)#ip address 192.168.12.1 255.255.255.0
ruijie(config-if)#ip ospf authentication-key key123
ruijie(config)#router ospf 1
ruijie(config-router)#network 192.168.12.0 0.0.0.255 area 0
ruijie(config-router)#area 0 authentication
```

例如：取消已经设置的区域 0 的 MD5 验证方式。

```
ruijie(config)#router ospf 1
ruijie(config-router)#network 192.168.12.0 0.0.0.255 area 0
ruijie(config-router)#no area 0 authentication
```

第 2 步：定义验证密码。

OSPF 协议同 RIP 一样，验证密码设置也分为明文验证和密码验证，分别描述如下。

（1）明文验证密码。OSPF 明文验证就是将验证密码以明文方式直接插入到 OSPF 报头中发送给邻居路由器。因密码不一致，两台直接连接的设备不能建立 OSPF 邻居关系，当然也不能进行路由信息的交换。不同接口的密码可以配置不一样，但所有连接在同一物理网段上的路由设备，必须配置一样的密码。

配置接口明文验证密码的命令格式如下：

设置接口验证密码命令

ruijie(config-if)# ip ospf authentication-key *key*

设备名称为ruijie
处于接口配置模式

验证密码字符串

其中，*key* 为密码，最多可以由 8 个字母或数字组成。

默认情况下，没有配置验证密码。使用该命令的 no 形式删除已配置的明文验证密码，如 ruijie(config-if)# no ip ospf authentication-key。

例如：某网络中，设置路由器 FastEthernet 0/0 接口的 OSPF 明文验证密码为 key123。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip address 172.16.10.1 255.255.255.0
ruijie(config-if)#ip ospf authentication-key key123
```

例如：删除已经配置路由器 FastEthernet 0/0 接口的 OSPF 明文验证密码 key123。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip address 172.16.10.1 255.255.255.0
ruijie(config-if)#no ip ospf authentication-key
```

（2）密文验证密码。OSPF 密文验证就是将验证密码以密文方式直接插入到 OSPF 报头中发送给邻居路由器。因密钥不一致，两台直接连接的设备是不能建立 OSPF 邻居关系，当然也不能进行路由信息的交换。不同接口的密码可以配置不一样，但所有连接在同一物理网段上的路由设备，必须配置一样的密码。邻居路由设备相同的密码标志对应的密码必须一样。

配置接口密文验证密码的命令格式如下：

其中，*key-id* 为密码标志符，从 1 到 255；*key* 为密码字符串，最多可以由 16 个字母或数字组成。

默认情况下，没有配置 MD5 密码。使用该命令的 no 形式删除已配置的 OSPF 报文 MD5 验证密码，如 ruijie(config-if)# no ip ospf message-digest-key。

MD5 验证密码修改遵循一个先加后删除的原则。当增加一台路由设备的 OSPF 的 MD5 验证密码时，该路由设备会认为其他路由设备还没有用新的密码，因此会分别用不同的密码发送多份 OSPF 报文，直到确认邻居已经配置了新的密码。等所有路由设备配置了新的密码后，就可以删除旧的密码了。

例如：某网络中，在路由器 FastEthernet 0/0 接口上增加了一个新的 OSPF 验证密码，密码号为 5，密码为 hello5。

```
ruijie(config)#interface FastEthernet 0/0
ruijie(config-if)#ip address 172.16.24.2 255.255.255.0
ruijie(config-if)#ip ospf authentication message-digest
ruijie(config-if)#ip ospf message-digest-key 10 md5 hello10
ruijie(config-if)#ip ospf message-digest-key 5 md5 hello5
```

等所有的邻居都增加了新的密码后，全部路由设备都需要删除旧的密码。

```
ruijie(config)#interface FastEthernet 0/0
ruijie(config-if)#no ip ospf message-digest-key 10 md5 hello10
```

8.3　扩展知识

1. OSPF 特殊区域

在前面章节我们已经知道，OSPF 有两种区域类型：骨干区域和非骨干区域，并且 OSPF 要求所有非骨干区域要与骨干区域连接。

除了这两种区域类型外，OSPF 还包括如下特殊区域。

末节区域（Stub Area）：该区域中不包含 ASBR，并且不会接收外部路由信息（5 类 LSA）。对于 Stub 区域，如果要到达外部 AS 的话，需要使用到达 Stub 区域的 ABR 的默认路由。Stub 区域的好处是可以减小链路状态数据库和路由表的规模。

绝对末节区域（Totally Stubby Area）：该区域对于 LSA 进入的限制更加严格。绝对末节区域不接收外部路由信息（5 类 LSA）和路由汇总信息（3 类 LSA、4 类 LSA）。绝对末节区域的好处是最小化链路状态数据库和路由表的规模。

次末节区域（Not-So-Stubby Area，NSSA）：该区域类似末节区域和绝对末节区域的优点，但是 NSSA 区域中可以包含 ASBR，即可以存在外部路由，但是外部路由在 NSSA 区域中是以一种特殊的 LSA 类型（7 类 LSA）进行通告的。

这三类末节区域的共同目的是将默认路由注入到区域中，从而防止外部 LSA 和汇总 LSA 扩散到区域中，末节区域和绝对末节区域不接受外部路由，区域具有以下特征。

第一，只有一个出口或有多个出口，但不要求选择最佳路径。

第二，必须将末节区域中所有的 OSPF 路由器都配置为末节路由器。

第三，区域不会被用于虚拟链路的中转区域。

第四，不是骨干区域（Area 0）。

（1）末节区域及其应用配置。当一个区域为 OSPF 路由域的末节区域时，则该区域不作为过渡区域，也不会向 OSPF 路由域注入外部路由。Stub 区域设备只能学到三种路由：Stub 区域内部路由、其他区域路由、Stub 区域边界路由器通告的默认路由，如图 8.23 所示。由于没有大量的外部路由，所以 Stub 区域设备的路由表将很小，可以节约设备的资源，因此 Stub 区域的设备可以为中低端设备。

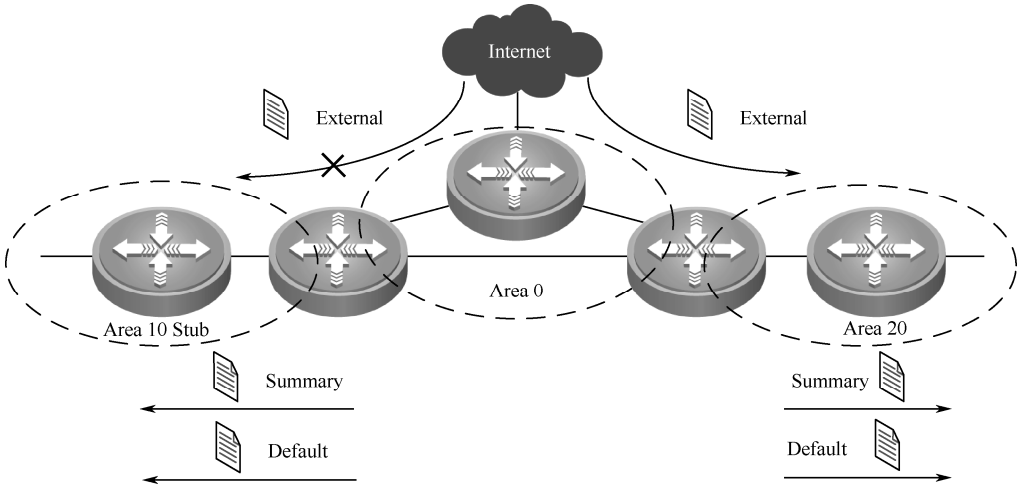


图 8.23　Stub 区域 LSA 类型

使用如下命令可以将区域配置为末节区域：

定义区域

ruijie(config-router)# area *area-id* stub

设备名称为ruijie
处于路由配置模式

末节区域

其中，*area-id* 为 Stub 区域的区域号。默认情况下没有定义末节区域。使用该命令的 no 形式删除末节区域定义，如 ruijie(config-router)#no area 10 stub，删除区域 10 的末节区域属性。配置完该命令后，ABR 将会为末节区域生成一条默认路由，并通告到末节区域。

OSPF 末节区域内所有的路由设备都必须执行 area stub 命令。ABR 只向末节区域发送三种链路状态描述（LSA）：类型 1（路由设备 LSA）、类型 2（网络 LSA）、类型 3（网络摘要 LSA）。从路由表角度看，末节区域中的路由设备只能学习到 OSPF 路由域内部的路由，包括由 ABR 产生的内部默认路由。末节区域的路由设备不能学习到 OSPF 路由域外部的路由。

例如：在如图 8.24 所示网络中，路由器 RA 中有 4 个子网 192.168.1.0、192.168.2.0、192.168.3.0、192.168.4.0，4 台路由器均启动 OSPF，Area 20 配置为 Stub 区域。

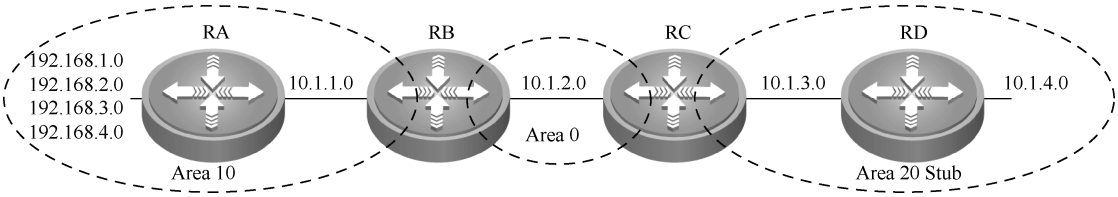


图 8.24　配置 Stub 区域

路由器 RA 配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 192.168.2.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 192.168.3.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop3
RA(config-if)#ip address 192.168.4.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#router ospf 1
RA(config-router)#network 192.168.1.0 0.0.0.255 area 10
RA(config-router)#network 192.168.2.0 0.0.0.255 area 10
RA(config-router)#network 192.168.3.0 0.0.0.255 area 10
RA(config-router)#network 192.168.4.0 0.0.0.255 area 10
RA(config-router)#network 10.1.1.0 0.0.0.255 area 10
RA(config-router)#end
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 10.1.1.2 255.255.255.0
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 10.1.2.1 255.255.255.0
RB(config-if)#exit
RB(config)#router ospf 1
RB(config-router)#network 10.1.1.0 0.0.0.255 area 10
RB(config-router)#network 10.1.2.0 0.0.0.255 area 0
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 10.1.2.2 255.255.255.0
RC(config-if)#exit
```

```
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 10.1.3.1 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 10.1.2.0 0.0.0.255 area 0
RC(config-router)#network 10.1.3.0 0.0.0.255 area 20
RC(config-router)#area 20 stub
RC(config-router)#end
RC#
```

路由器 RD 配置代码如下：

```
RD#configure terminal
RD(config)#interface fastethernet 0/0
RD(config-if)#ip address 10.1.3.2 255.255.255.0
RD(config-if)#exit
RD(config)#interface fastethernet 0/1
RD(config-if)#ip address 10.1.4.1 255.255.255.0
RD(config-if)#exit
RD(config)#router ospf 1
RD(config-router)#network 10.1.3.0 0.0.0.255 area 20
RD(config-router)#network 10.1.4.0 0.0.0.255 area 20
RD(config-router)#area 20 stub
RD(config-router)#end
RD#
```

通过 show ip route 命令查看路由器 RD 的路由表，发现 RD 没有学习到外部路由，但是收到了一条区域间默认路由及多条区域间路由，默认路由使用 O *IA 标志，区域间路由使用 O IA 标志。

（2）绝对末节区域及其应用配置。绝对末节区域进一步减少了链路状态数据库和路由表中的条目，提高了 OSPF 互联网络的稳定性和可扩展性。外部 LSA 及 3 类 LSA、4 类 LSA 都不能传播到绝对末节区域中，如图 8.25 所示。绝对末节区域中只有区域内路由和默认路由。

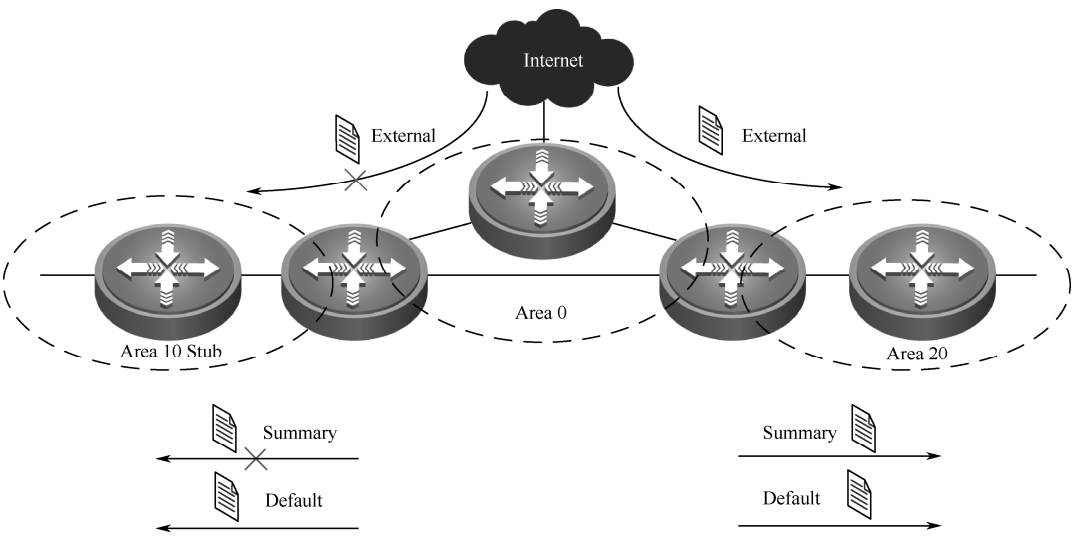


图 8.25　绝对末节区域 LSA 类型

使用如下命令可以将区域配置为绝对末节区域：

定义区域　　*禁止通告汇总路由*

ruijie(config-router)# area *area-id* **stub no-summary**

设备名称为ruijie
处于路由配置模式　　*末节区域*

其中，*area-id* 为 Stub 区域的区域号。默认情况下没有定义绝对末节区域。使用该命令的 no 形式删除绝对末节区域定义，如 ruijie(config-router)#no area 10 stub no-summary，删除区域 10 的绝对末节区域属性。配置完该命令后，ABR 将会为绝对末节区域生成一条默认路由，并通告到绝对末节区域。

在配置绝对末节区域时，只需在 ABR 上执行 area stub no-summary 命令，其他路由器执行不带 no-summary 参数的 area stub 命令。绝对末节区域的路由设备只能学到本区域内的路由，以及 ABR 产生的内部默认路由。

例如：在如图 8.26 所示网络中，路由器 RA 中有 4 个子网 192.168.1.0、192.168.2.0、192.168.3.0、192.168.4.0，4 台路由器均启动 OSPF，Area 20 配置为绝对末节区域。

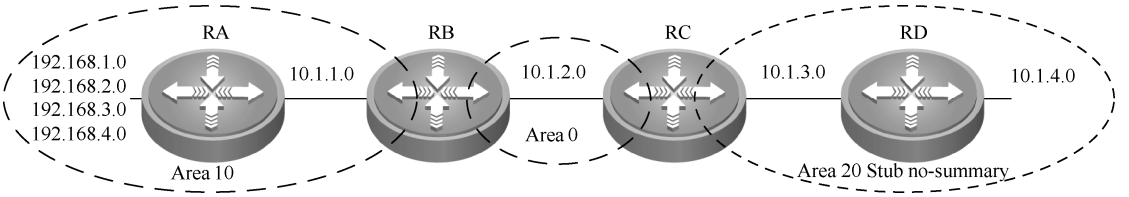


图 8.26　配置绝对末节区域

路由器 RA 配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 192.168.2.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 192.168.3.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop3
RA(config-if)#ip address 192.168.4.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#router ospf 1
RA(config-router)#network 192.168.1.0 0.0.0.255 area 10
RA(config-router)#network 192.168.2.0 0.0.0.255 area 10
RA(config-router)#network 192.168.3.0 0.0.0.255 area 10
```

```
RA(config-router)#network 192.168.4.0 0.0.0.255 area 10
RA(config-router)#network 10.1.1.0 0.0.0.255 area 10
RA(config-router)#end
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 10.1.1.2 255.255.255.0
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 10.1.2.1 255.255.255.0
RB(config-if)#exit
RB(config)#router ospf 1
RB(config-router)#network 10.1.1.0 0.0.0.255 area 10
RB(config-router)#network 10.1.2.0 0.0.0.255 area 0
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 10.1.2.2 255.255.255.0
RC(config-if)#exit
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 10.1.3.1 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 10.1.2.0 0.0.0.255 area 0
RC(config-router)#network 10.1.3.0 0.0.0.255 area 20
RC(config-router)#area 20 stub no-summary
RC(config-router)#end
RC#
```

路由器 RD 配置代码如下：

```
RD#configure terminal
RD(config)#interface fastethernet 0/0
RD(config-if)#ip address 10.1.3.2 255.255.255.0
RD(config-if)#exit
RD(config)#interface fastethernet 0/1
RD(config-if)#ip address 10.1.4.1 255.255.255.0
RD(config-if)#exit
RD(config)#router ospf 1
RD(config-router)#network 10.1.3.0 0.0.0.255 area 20
RD(config-router)#network 10.1.4.0 0.0.0.255 area 20
RD(config-router)#area 20 stub
```

```
RD(config-router)#end
RD#
```

通过 show ip route 命令查看路由器 RD 的路由表，发现 RD 没有学习到外部路由，但是只收到了一条区域间默认路由，默认路由使用 O *IA 标志。没有收到区域间路由。

（3）次末节区域及其应用配置。次末节区域 NSSA 是对 OSPF Stub 区域的一个扩展，NSSA 也通过阻止第 5 类 LSA（AS-external- LSA）向 NSSA 的泛洪来降低设备的资源的消耗，但是和 Stub 区域不同的是 NSSA 可以注入一定数量的自治系统外部的路由信息到 OSPF 的路由选择域。

通过重分发，NSSA 允许输入类型 7 的自治系统外部路由到 NSSA 区域，这些类型 7 的外部 LSA 在 NSSA 的区域边界设备上将被转换成类型 5 的 LSA 并且泛洪到整个自治系统，在转换过程中可以实现对外部路由的汇总和过滤，如图 8.27 所示。

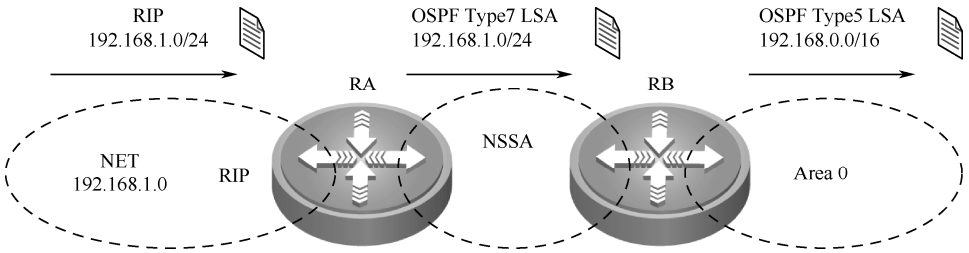


图 8.27　NSSA 区域

配置 NSSA 时要注意以下几点。

骨干区域不能配置成 NSSA，不能将 NSSA 作为虚拟连接的传输区域。

如果想将一个区域配置成 NSSA，所有连接到 NSSA 区域的设备必须使用 area nssa 命令将该区域配置成 NSSA 属性。

要配置区域成为一个 NSSA 区域，在路由进程配置模式中执行以下命令：

```
ruijie(config-router)#area area-id nssa [ no-redistribution] [default-information-originate] [no-summary]
```

定义区域　*禁止重分发路由进入NSSA*　*禁止汇总路由进入NSSA*

设备名称为ruijie 处于路由配置模式　*次末节区域*　*产生默认路由*

其中，*area-id* 为 NSSA 区域的区域号；no-redistribution 为当该路由器是一个 NSSA ABR 时，如果不想将重分发的路由信息导入 NSSA 区域时，可以使用该选项；default-information-originate 为产生默认的 Type 7 LSA 并通告 NSSA 区域，该选项只在 NSSA ABR 或者 ASBR 上有效；no-summary 为阻止 NSSA 的边界路由设备（ABR）发送第三和第四类 LSA 进入 NSSA 区域。

默认情况下，路由设备没有定义 NSSA 区域。该命令的 no 形式删除 NSSA 区域或 NSSA 区域的配置，如 ruijie(config-router)# no area 10 nssa no-summary，删除 Area 10 的 NSSA 属性。

例如：在某网络中区域 10 配置为 NSSA，在通过路由器 RB 导入 RA 的 RIP 路由信息。在路由器 RC 上配置 NSSA 的 No-Summary 功能，禁止外部 LSA 和汇总（3 类、4 类）LSA 进入 NSSA 区域，向 NSSA 区域通告一条默认路由信息，如图 8.28 所示。

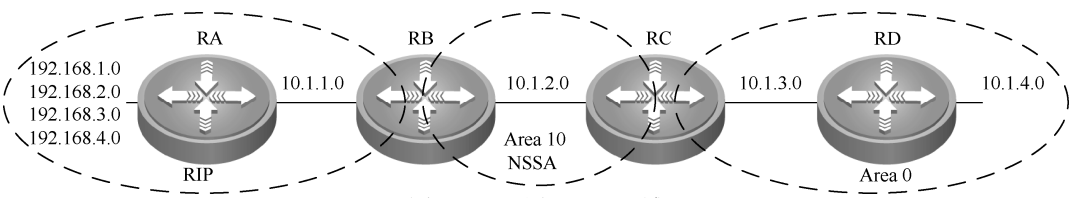


图 8.28　配置 NSSA 区域

路由器 RA 配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 192.168.2.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 192.168.3.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop3
RA(config-if)#ip address 192.168.4.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#router rip
RA(config-router)#version 2
RA(config-router)#network 192.168.1.0
RA(config-router)#network 192.168.2.0
RA(config-router)#network 192.168.3.0
RA(config-router)#network 192.168.4.0
RA(config-router)#network 10.0.0.0
RA(config-router)#no auto-summary
RA(config-router)#end
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 10.1.1.2 255.255.255.0
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 10.1.2.1 255.255.255.0
RB(config-if)#exit
RB(config)#router rip
RB(config-router)#version 2
RB(config-router)#network 10.0.0.0
RB(config-router)#no auto-summary
RB(config-router)#redistribute connected
RB(config-router)#redistribute ospf 10 metric 1
RB(config-router)#exit
RB(config)#router ospf 10
```

```
RB(config-router)#network 10.1.2.0 0.0.0.255 area 10
RB(config-router)#redistribute connected
RB(config-router)#redistribute rip metric 50 subnets
RB(config-router)#area 10 nssa
RB(config-router)#default-metric 50
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 10.1.2.2 255.255.255.0
RC(config-if)#exit
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 10.1.3.1 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 10.1.2.0 0.0.0.255 area 10
RC(config-router)#network 10.1.3.0 0.0.0.255 area 0
RC(config-router)#area 10 nssa no-summary
RC(config-router)#end
RC#
```

路由器 RD 配置代码如下：

```
RD#configure terminal
RD(config)#interface fastethernet 0/0
RD(config-if)#ip address 10.1.3.2 255.255.255.0
RD(config-if)#exit
RD(config)#interface fastethernet 0/1
RD(config-if)#ip address 10.1.4.1 255.255.255.0
RD(config-if)#exit
RD(config)#router ospf 1
RD(config-router)#network 10.1.3.0 0.0.0.255 area 0
RD(config-router)#network 10.1.4.0 0.0.0.255 area 0
RD(config-router)#end
RD#
```

（4）OSPF 默认路由。通过前面的学习已经了解到 ABR 能够对末节区域、绝对末节区域、次末节区域自动生成一条默认路由并且通告到相应区域中。但是对于标准区域，路由器是不会自动生成默认路由信息的。如果要让 ASBR 路由器自动生成一条默认路由信息并通告至标准区域，必须执行如下命令：

其中，*always* 表示无论本地路由器是否存在默认路由都通告默认路由信息；*metric-value* 为指定默

认路由的度量值，默认值为 1；*type-value* 为指定外部路由度量值的类型（E1 或 E2 类型），默认值为 E2 类型。

在该命令上也可以附加[route-map map-name]，以通过 route-map 限制条件。

默认情况下不产生默认路由。使用该命令的 no 形式关闭配置的默认路由，如 ruijie(config-router)# no default-information originate always metric 40。

例如：在某网络中，要求 OSPF 自治系统边界路由器产生一条外部默认路由，注入到 OSPF 路由域中，类型设置为 1，量度值设置为 50。

```
ruijie(config)#router ospf 1
ruijie(config-router)#network 172.16.24.0 0.0.0.255 area 0
ruijie(config-router)#default-information originate always metric 50 metric-type 1
```

2. OSPF 虚拟链路及其应用配置

在 OSPF 路由域设计中，要求非骨干区域必须与骨干区域连接，非骨干区域之间通过骨干区域交换路由更新报文。

如果由于某种原因骨干区域内连接断开，不能实现骨干区域内直接连接，可以在 ABR 之间建立虚拟链路将骨干区域接续起来，否则网络通信将出现问题，如图 8.29 所示。

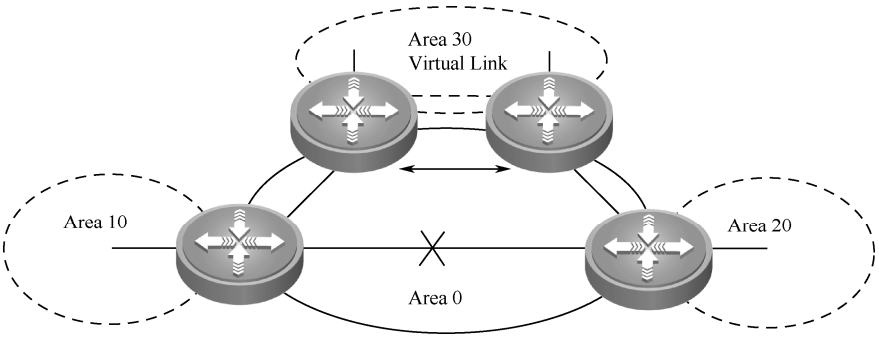


图 8.29　骨干区域断接

如果由于网络拓扑结构的限制无法保证非骨干区域直接通过物理连接到骨干区域上，也可以建立虚拟链路实现与骨干区域的逻辑连接，如图 8.30 所示。

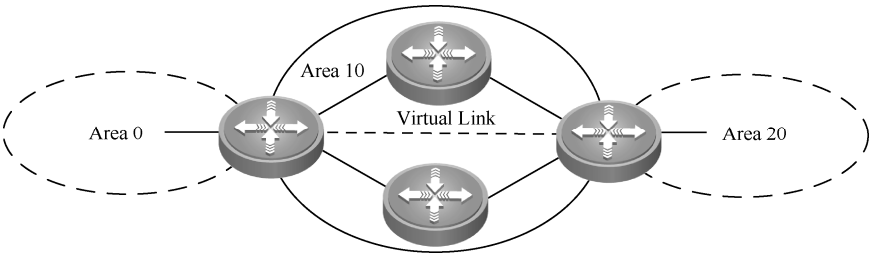


图 8.30　非骨干区域断接

虚拟链路需要在两个区域边界设备（ABR）之间创建，两个 ABR 共同所属的区域成为过渡区域。末节区域（Stub Area）和 NSSA 区域是不能作为过渡区域的。虚拟连接可以看成是在两台 ABR 之间通过传输区域建立的一条逻辑上的连接通道，它的两端必须是 ABR，而且必须在两端同时配置方可生效。虚拟链路由对端设备的 Router-ID 号来标志。

使用如下命令创建虚拟链路：

定义区域

ruijie(config-router)#area *area-id* **virtual-link** *router-id*

设备名称为ruijie
处于路由配置模式

设置虚拟链路另一端Router-ID

其中，*area-id* 为 OSPF 过渡区域号，区域号可以是一个十进制整数值，也可以是一个 IP 地址；*router-id* 为虚拟链路另一端路由器的 Router-ID。

使用该命令的 no 形式删除已定义的虚拟链路，如 ruijie(config-router)#no area 10 virtual-link 1.1.1.1 删除在区域 10 中的与 Rouer-ID 为 1.1.1.1 的虚拟链路。

例如：公司网络中由于地理位置原因，区域 10 不能与区域 0 直接物理连接，需要经过区域 20 连接到骨干区域 0。因此，需要在区域 20 中的两台 ABR 之间直接创建一条虚拟链路，在路由器 RB 上指定 RC 的 Router-ID，在路由器 RC 上指定 RB 的 Router-ID，如图 8.31 所示。

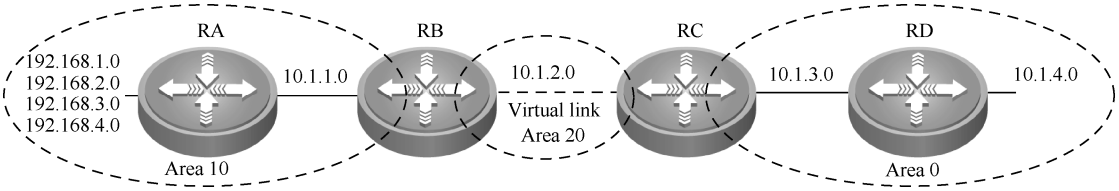


图 8.31　配置虚拟链路

路由器 RA 配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 192.168.2.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 192.168.3.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop3
RA(config-if)#ip address 192.168.4.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#router ospf 1
RA(config-router)#router-id 1.1.1.1
RA(config-router)#network 192.168.1.0 0.0.0.255 area 10
RA(config-router)#network 192.168.2.0 0.0.0.255 area 10
RA(config-router)#network 192.168.3.0 0.0.0.255 area 10
```

```
RA(config-router)#network 192.168.4.0 0.0.0.255 area 10
RA(config-router)#network 10.1.1.0 0.0.0.255 area 10
RA(config-router)#end
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 10.1.1.2 255.255.255.0
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 10.1.2.1 255.255.255.0
RB(config-if)#exit
RB(config)#router ospf 1
RB(config-router)#router-id 2.2.2.2
RB(config-router)#network 10.1.1.0 0.0.0.255 area 10
RB(config-router)#network 10.1.2.0 0.0.0.255 area 20
RB(config-router)#area 20 virtual-link 3.3.3.3
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 10.1.2.2 255.255.255.0
RC(config-if)#exit
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 10.1.3.1 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#router-id 3.3.3.3
RC(config-router)#network 10.1.2.0 0.0.0.255 area 20
RC(config-router)#network 10.1.3.0 0.0.0.255 area 0
RC(config-router)#area 20 virtual-link 2.2.2.2
RC(config-router)#end
RC#
```

路由器 RD 配置代码如下：

```
RD#configure terminal
RD(config)#interface fastethernet 0/0
RD(config-if)#ip address 10.1.3.2 255.255.255.0
RD(config-if)#exit
RD(config)#interface fastethernet 0/1
RD(config-if)#ip address 10.1.4.1 255.255.255.0
RD(config-if)#exit
RD(config)#router ospf 1
```

```
RD(config-router)#router-id 4.4.4.4
RD(config-router)#network 10.1.3.0 0.0.0.255 area 0
RD(config-router)#network 10.1.4.0 0.0.0.255 area 0
RD(config-router)#end
RD#
```

8.4　演示实例

1. 任务描述

在某网络中，OSPF 区域 1 包含路由器 RA 和 RB，OSPF 区域 0 包含路由器 RB 和 RC，RB 为区域边界路由器。路由器 RA 中有 172.16.10.0、172.16.11.0、172.16.12.0 三个子网，可以在路由器 RB 上通过路由汇总将区域 1 中的三条路由条目汇总成一条路由通告到区域 0 中，如图 8.32 所示。

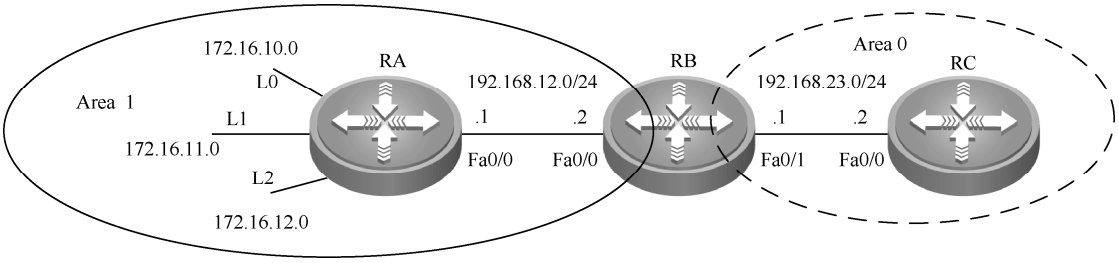


图 8.32　区域间路由汇总

2. 操作步骤

路由器 RA 配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 172.16.10.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 172.16.11.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 172.16.12.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#exit
RA(config)#router ospf 1
RA(config-router)#network 172.16.10.0 0.0.15.255 area 1
RA(config-router)#network 172.16.11.0 0.0.15.255 area 1
RA(config-router)#network 172.16.12.0 0.0.15.255 area 1
RA(config-router)#network 192.168.12.0 0.0.0.255 area 1
RA(config-router)#end
```

```
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 192.168.12.2 255.255.255.0
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 192.168.23.1 255.255.255.0
RB(config-if)#exit
RB(config)#router ospf 1
RB(config-router)#network 192.168.23.0 0.0.0.255 area 0
RB(config-router)#network 192.168.12.0 0.0.0.255 area 1
RB(config-router)#area 1 range 172.16.0.0 255.255.240.0
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 192.168.23.2 255.255.255.0
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 192.168.23.0 0.0.0.255 area 0
RC(config-router)#end
RC#
```

在区域0的路由器RC上执行show ip route 命令显示其路由表，验证是否已经将区域1中172.16.0.0的三个子网路由汇总成一条汇总路由并通告至区域0中。

8.5 训练任务

1. 背景描述

某网络中运行 OSPF 路由协议，为了保证网络中路由信息交换的安全性，需要控制网络中的路由更新只在可信任的路由器之间进行，网络拓扑如图 8.33 所示。

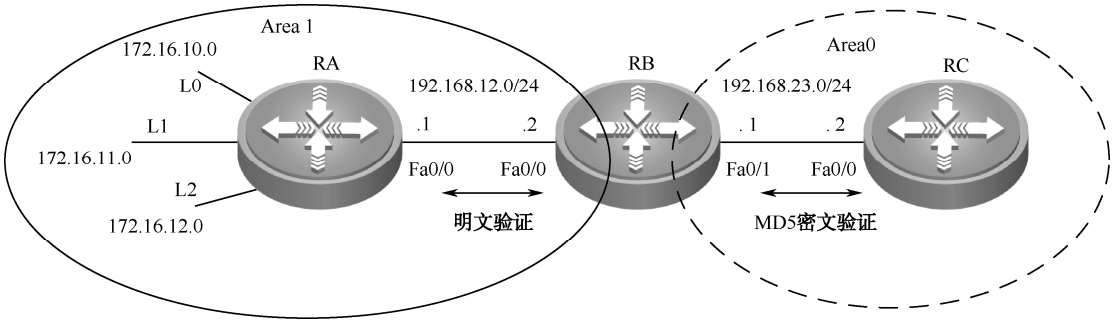


图 8.33 OSPF 路由验证

在路由器 RA 与路由器 RB 之间采用明文验证方式；在路由器 RB 与路由器 RC 之间采用 MD5 密文验证方式。

2. 操作提示

在路由器 RA 上明文验证密码为 key，配置代码如下：

```
RA#configure terminal
RA(config)#interface loop0
RA(config-if)#ip address 172.16.10.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop1
RA(config-if)#ip address 172.16.11.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loop2
RA(config-if)#ip address 172.16.12.1 255.255.255.0
RA(config-if)#ip ospf authenticaion
RA(config-if)#ip ospf authentication-key key123
RA(config-if)#exit
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#exit
RA(config)#router ospf 1
RA(config-router)#network 172.16.10.0 0.0.15.255 area 1
RA(config-router)#network 172.16.11.0 0.0.15.255 area 1
RA(config-router)#network 172.16.12.0 0.0.15.255 area 1
RA(config-router)#network 192.168.12.0 0.0.0.255 area 1
RA(config-router)#end
RA#
```

路由器 RB 配置代码如下：

```
RB#configure terminal
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 192.168.12.2 255.255.255.0
RB(config-if)#ip ospf authenticaion
RB(config-if)#ip ospf authentication-key key123
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 192.168.23.1 255.255.255.0
RB(config-if)#ip ospf authentication message-digest
RB(config-if)#ip ospf message-digest-key 1 md5 key567
RB(config-if)#exit
RB(config)#router ospf 1
RB(config-router)#network 192.168.23.0 0.0.0.255 area 0
RB(config-router)#network 192.168.12.0 0.0.0.255 area 1
RB(config-router)#area 1 range 172.16.0.0 255.255.240.0
RB(config-router)#end
```

```
RB#
```

路由器 RC 配置代码如下：

```
RC#configure terminal
RC(config)#interface fastethernet 0/0
RC(config-if)#ip address 192.168.23.2 255.255.255.0
RC(config-if)#ip ospf authentication message-digest
RC(config-if)#ip ospf message-digest-key 1 md5 key567
RC(config-if)#exit
RC(config)#router ospf 1
RC(config-router)#network 192.168.23.0 0.0.0.255 area 0
RC(config-router)#end
RC#
```

在通过修改密码后，利用 show ip ospf neighber 命令查看邻居状态，验证 OSPF 验证效果。

练习题

1. 选择题

（1）下面关于 OSPF 路由协议的描述错误的是（　　）。

A．OSPF 路由协议是一种链路状态路由选择协议，是内部网关协议 IGP 之一

B．OSPF 路由协议将自己的路由表发送给相邻路由器

C．OSPF 路由协议只有当网络链路状态发送变化时，才发送更新报文

D．OSPF 路由协议是分区的路由协议

（2）下面关于 OSPF 路由协议优点的描述错误的是（　　）。

A．由于划分区域，每个区域内的路由器数量减少，路由信息流量也就减少了

B．由于区域内链路减小了，因此链路状态数据库数据量也减小了

C．同一区域中的每台路由器都仅为该区域计算其 SPF 树，路由表减小了

D．每台路由器 LSA 在整个 OSPF 区域内传播，更新报文数量更大了

（3）下面不属于 OSPF 协议分组的是（　　）。

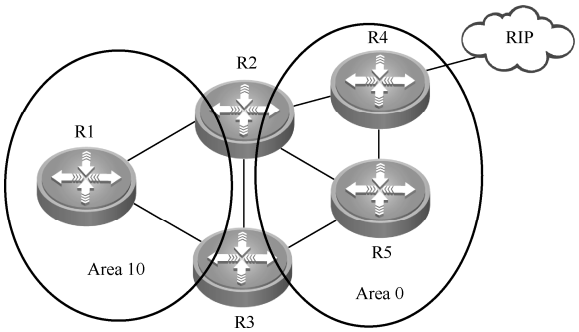
A．问候分组（Hello）

B．数据库描述分组（Database Description）

C．链路故障分组（Link Failure）

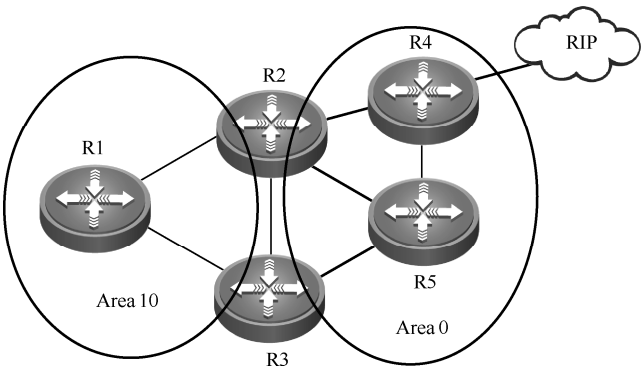
D．链路状态更新分组（Link-State Update）

（4）在下图所示网络中，根据路由器所处位置，确定路由器角色，R3 是（　　）。



A．区域内路由器　B．区域边界路由器　C．骨干路由器　D．自治系统边界路由器

（5）在下图所示网络中，根据路由器所处位置，确定路由器角色，R4 是（　　）。



A．区域内路由器　B．区域边界路由器　C．骨干路由器　D．自治系统边界路由器

2. 简答题

（1）描述 OSPF 协议工作过程。

（2）OSPF 网络划分区域后，网络中路由器根据位置不同其扮演的角色也不同。说明 OSPF 路由器有哪些类型。

（3）说明 OSPF 协议将自治系统划分为不同区域的优点是什么。

项目 9　路由选择控制与过滤

通过前面章节的学习，能够配置动态路由协议，使网络中路由器学习到其他网段路由。但是，考虑到网络安全因素，往往需要对动态路由更新信息加以限制。限制路由更新有多种方法，包括被动接口技术、分发列表技术及调整 AD 值等。本章介绍被动接口技术、分发列表技术及调整 AD 值等相关知识及配置技能，达到对路由信息进行控制和过滤的目的。

知识点、技能点

1. 了解路由器被动接口、分发列表及调整 AD 值的功能及应用;
2. 掌握路由器被动接口、分发列表及调整 AD 值的配置方法;
3. 掌握网络工程调试方法及技能。

9.1　问题提出

在网络应用中，为了防止其他路由器动态地学习到设备路由信息，需要禁止路由器接口对外发布路由更新信息或过滤路由信息，为此，路由器被动接口技术、分发列表技术及调整 AD 值技术被广泛应用。

9.2　相关知识

9.2.1　被动接口

1. 被动接口概述

为了防止其他路由器动态地学习到本路由器的路由信息，可以配置本路由器的接口为被动接口 passive-interface，不允许路由更新报文从该路由器接口发送出去。既然没有路由更新报文从路由器接口发送出去，该接口所连接的路由器自然就学不到路由信息了，该特性可以应用在所有的 IGP 路由协议上。

被动接口将阻止通过该接口发送指定协议的路由选择更新。被动接口不参与路由进程中，在 OSPF 路由域中，如果路由器接口配置了 passive-interface，该接口在 OSPF 路由域中就表现为一个孤立网络，而且该接口将不发送更新路由信息也不接收更新路由信息。实际上，配置了 passive-interface，该接口将不发送 OSPF 的 Hello 报文，所以该接口就不可能有邻居的存在，也不会交换路由。

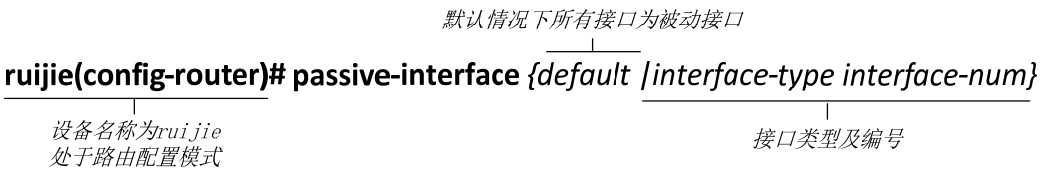
RIP 的表现与 OSPF 不太一样，RIP 路由器接口配置了 passive-interface 后，不发送路由更新报文，但是还可以接收路由更新报文，而且 RIP 可以通过定义邻居的方式只给指定的邻居发送更新报文。

2. 配置命令

（1）设置被动接口命令。路由器配置命令 passive-interface type number [default]用于禁止通过指定的路由器接口发送路由选择更新，该命令可用于将特定接口设置为被动状态，也可以将所有路由器接

口设置为被动状态，使用 default 选项将所有路由器接口设置为被动状态。

该命令格式如下：



其中，*default* 为将路由器所有接口默认状态设置为被动状态；*interface-type* 为指定被动接口类型；*interface-num* 为指定被动接口编号。

默认情况下，接口为非 passive 模式。该命令的 no 形式重新启用发送更新报文的功能，如 ruijie(config-router)#no passive-interface Fastethernet 0/0 命令将 Fastethernet 0/0 接口恢复为默认值，即非 passive 模式。

例如：设置接口 Fastethernet 0/1 为 passive 模式。

```
ruijie(config)#router rip
ruijie(config-router)#passive-interface fastethernet 0/1
ruijie(config-router)#
```

例如：设置所有接口为 passive 模式，然后设置 Fastethernet 0/0 为非 passive 模式。

```
ruijie(config-router)#passive-interface default
ruijie(config-router)#no passive-interface fastethernet 0/0
```

（2）指定邻居 IP 地址。RIPv1 默认使用 IP 广播地址（255.255.255.255）通告路由信息，RIPv2 默认使用组播地址（224.0.0.9）通告路由信息。如果不希望广播网或非广播多路访问网上的全部设备，均可接收到路由信息，可以用路由进程配置命令 passive-interface 将相应接口设置为被动接口，然后只定义某些邻居可以接收到路由信息。该命令不会影响 RIP 信息报文的接收。设置了 passive 属性的接口，重启后也不会发送请求报文。该命令格式如下：

其中，*ip-address* 为邻居的 IP 地址。应该是本地设备直连网络地址。默认情况下没有定义邻居，该命令的 no 形式删除邻居定义，如 ruijie(config-router)# no neighbor。

例如：R0 路由器 Fastethernet 0/0 接口连接路由器 R1 和 R2，R1 接口地址为 172.16.1.2/24，R2 接口地址为 172.16.2.2/24，要求 R0 路由器只将路由更新信息发送给 R1。

```
R0(config)#router rip
R0(config-router)#passive-interface Fastethernet 0/0
R0(config-router)#neighbor 172.16.1.2
```

例如：在如图 9.1 所示网络中，路由器 RA 接口 Fa0/0 和三层交换机 SW 接口 Fa0/1 连接，交换机通过接口 Fa0/24 与服务器区服务器连接。交换机及路由器均启动动态路由协议 RIPv2，路由更新报文从所有网络接口发送，这样所有服务器也将收到路由更新报文，但服务器并不需要路由更新报文。为了减少路由更新报文对服务器的影响，将交换机接口 Fa0/24 接口配置为 passive-interface。

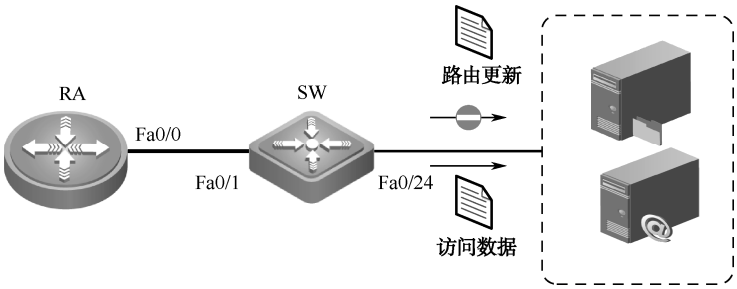


图 9.1　passive-interface 应用

设备参数配置如下。

第 1 步：基本配置。

```
ruijie(config)#hostname RA
RA(config)#interface interface fastethernet  0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface interface loopback 0
RA(config-if)#ip address 10.10.1.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#router rip
RA(config-router)#version 2
RA(config-router)#network 10.0.0.0
RA(config-router)#network 192.168.12.0
RA(config-router)#no auto-summary
RA(config-router)#end
RA#
ruijie(config)#hostname SW
SW(config)#interface FastEthernet 0/1
SW(config-if)#no switchport
SW(config-if)#ip address 192.168.12.2 255.255.255.0
SW(config-if)#exit
SW(config)#interface FastEthernet 0/24
SW(config-if)#no switchport
SW(config-if)#ip address 172.16.1.1 255.255.255.0
SW(config-if)#exit
SW(config)#
SW(config)#router rip
SW(config-router)#version 2
SW(config-router)#network 172.16.0.0
SW(config-router)#network 192.168.12.0
SW(config-router)#no auto-summary
SW(config-router)#end
SW#
```

第 2 步：观察交换机 Fa0/24 接口发送 RIP 更新报文的情况。

```
SW#debug ip rip packet send
*Jun    6 23:36:27: %7:    [RIP] Output timer expired to send reponse
*Jun    6 23:36:27: %7:    [RIP] Prepare to send MULTICAST response...
*Jun    6 23:36:27: %7:    [RIP] Building update entries on FastEthernet 0/1
*Jun    6 23:36:27: %7:          172.16.1.0/24 via 0.0.0.0 metric 1 tag 0
*Jun    6 23:36:27: %7:    [RIP] Send packet to 224.0.0.9 Port 520 on FastEthernet 0/1
*Jun    6 23:36:27: %7:    [RIP] Prepare to send MULTICAST response...
*Jun    6 23:36:27: %7:    [RIP] Building update entries on FastEthernet 0/24
*Jun    6 23:36:27: %7:          10.1.1.0/24 via 0.0.0.0 metric 2 tag 0
*Jun    6 23:36:27: %7:          192.168.12.0/24 via 0.0.0.0 metric 1 tag 0
*Jun    6 23:36:27: %7:    [RIP] Send packet to 224.0.0.9 Port 520 on FastEthernet 0/24
```

从捕获记录看，交换机从接口 Fa0/1 及 Fa0/24 均发送路由更新报文。

第 3 步：配置被动接口。

```
SW(config)#router rip
SW(config-router)#passive-interface FastEthernet 0/0
SW(config-router)#end
SW#
```

第 4 步：再次观察交换机 Fa0/24 发送 RIP 更新报文的情况。

```
SW#debug ip rip packet send
 *Jun    6 23:38:27: %7:    [RIP] Output timer expired to send reponse
 *Jun    6 23:38:27: %7:    [RIP] Prepare to send MULTICAST response...
 *Jun    6 23:38:27: %7:    [RIP] Building update entries on FastEthernet 0/1
 *Jun    6 23:38:27: %7:          172.16.1.0/24 via 0.0.0.0 metric 1 tag 0
 *Jun    6 23:38:27: %7:    [RIP] Send packet to 224.0.0.9 Port 520 on FastEthernet 0/1
 *Jun    6 23:38:57: %7:    [RIP] Output timer expired to send reponse
 *Jun    6 23:38:57: %7:    [RIP] Prepare to send MULTICAST response...
 *Jun    6 23:38:57: %7:    [RIP] Building update entries on FastEthernet 0/1
 *Jun    6 23:38:57: %7:          172.16.1.0/24 via 0.0.0.0 metric 1 tag 0
 *Jun    6 23:38:57: %7:    [RIP] Send packet to 224.0.0.9 Port 520 on FastEthernet 0/1
```

从捕获记录看，交换机只从接口 Fa0/1 发送路由更新报文，Fa0/24 接口已禁止发送更新报文了。

注意：对于 RIP 路由协议，阻止了路由更新报文发送，但是可以学习到 RIP 路由更新报文。对于 OSPF 路由协议，被动接口阻止了 Hello 包，无法建立邻居关系，同样，无法交换路由信息。

3. 应用案例

（1）背景描述。

某公司局域网内两个部门通过两台交换机 Switch1 和 Switch2 与路由器相连，并在 Router1、Switch1 及 Switch2 上配置动态路由协议 RIPv2。考虑到安全等原因，局域网路由器 Router1 的 RIP 路由更新信息只允许发送到交换机 Switch1，其他交换机不允许接收到路由更新信息，网络拓扑结构如图 9.2 所示。

在路由器 Router1 的接口 Fa0/0 上配置 passive-interface 状态，限制 RIP 更新信息从 Fa0/0 接口发送出去。但是，考虑到允许交换机 Switch1 接收 RIP 更新信息，需要在 Fa0/0 上配置单播报文发送配置，使得路由器 Router1 从 Fa0/0 接口发送单播报文到 Switch1 交换机，实现路由信息更新功能。

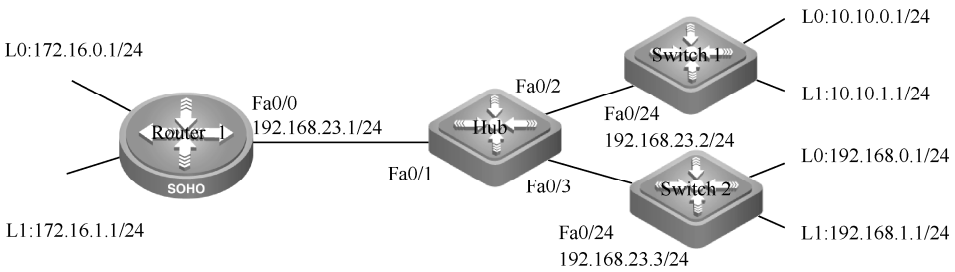


图 9.2　被动接口与单播更新

提示：*为了做实验方便，可以将上述图中 Hub 用交换机代替，交换机保持在默认状态，无须任何配置。*

（2）操作步骤。

第 1 步：配置路由器 Router1。

```
ruijie(config)#hostname Router1
Router1(config)#interface Fastethernet 0/0
Router1(config-if)#ip address 192.168.23.1 255.255.255.0
Router1(config-if)#no shutdown
Router1(config-if)#exit
Router1(config)#interface Loopback 0
Router1(config-if)#ip address 172.16.0.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#interface Loopback 1
Router1(config-if)#ip address 172.16.1.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#router rip
Router1(config-router)#version 2
Router1(config-router)#passive-interface FastEthernet 0/0
Router1(config-router)#network 172.16.0.0
Router1(config-router)#network 192.168.12.0
Router1(config-router)#neighbor 192.168.23.2
Router1(config-router)#no auto-summary
Router1(config-router)#timers basic 5 10 15
Router1(config-router)#end
Router1#
```

第 2 步：配置三层交换机 Switch1。

```
ruijie(config)#hostname Switch1
Switch1(config)#interface FastEthernet 0/24
Switch1(config-if)#no switchport
Switch1(config-if)#ip address 192.168.23.2 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#interface Loopback 0
Switch1(config-if)#ip address 10.10.0.1 255.255.255.0
Switch1(config-if)#exit
```

```
Switch1(config)#interface Loopback 1
Switch1(config-if)#ip address 10.10.1.1 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#router rip
Switch1(config-router)#version 2
Switch1(config-router)#network 10.0.0.0
Switch1(config-router)#network 192.168.23.0
Switch1(config-router)#no auto-summary
Switch1(config-router)#timers basic 5 10 15
Switch1(config-router)#end
Switch1#
```

第 3 步：配置三层交换机 Switch2。

```
ruijie(config)#hostname Switch2
Switch2(config)#interface FastEthernet 0/24
Switch2(config-if)#no switchport
Switch2(config-if)#ip address 192.168.23.3 255.255.255.0
Switch2(config-if)#exit
Switch2(config)#interface Loopback 0
Switch2(config-if)#ip address 192.168.0.1 255.255.255.0
Switch2(config-if)#exit
Switch2(config)#interface Loopback 1
Switch2(config-if)#ip address 192.168.1.1 255.255.255.0
Switch2(config-if)#exit
Switch2(config)#router rip
Switch2(config-router)#version 2
Switch2(config-router)#network 192.168.0.0
Switch2(config-router)#network 192.168.1.0
Switch2(config-router)#network 192.168.23.0
Switch2(config-router)#no auto-summary
Switch2(config-router)#timers basic 5 10 15
Switch2(config-router)#end
Switch2#
```

第 4 步：查看路由表。

```
Router1#show ip route
R     10.10.0.0/24 [120/1] via 192.168.23.2, 00:07:07, FastEthernet 0/0
R     10.10.1.0/24 [120/1] via 192.168.23.2, 00:07:07, FastEthernet 0/0
C     172.16.0.0/24 is directly connected, Loopback 0
C     172.16.0.1/32 is local host.
C     172.16.1.0/24 is directly connected, Loopback 1
C     172.16.1.1/32 is local host.
R     192.168.0.0/24 [120/1] via 192.168.23.3, 00:09:33, FastEthernet 0/0
R     192.168.1.0/24 [120/1] via 192.168.23.3, 00:09:33, FastEthernet 0/0
C     192.168.23.0/24 is directly connected, FastEthernet 0/0
```

```
C       192.168.23.1/32 is local host.
```

```
Switch1#show ip route
C       10.10.0.0/24 is directly connected, Loopback 0
C       10.10.0.1/32 is local host.
C       10.10.1.0/24 is directly connected, Loopback 1
C       10.10.1.1/32 is local host.
R       172.16.0.0/24 [120/1] via 192.168.23.1, 00:24:26, FastEthernet 0/24
R       172.16.1.0/24 [120/1] via 192.168.23.1, 00:24:26, FastEthernet 0/24
R       192.168.0.0/24 [120/1] via 192.168.23.3, 00:29:57, FastEthernet 0/24
R       192.168.1.0/24 [120/1] via 192.168.23.3, 00:29:57, FastEthernet 0/24
C       192.168.23.0/24 is directly connected, FastEthernet 0/24
C       192.168.23.2/32 is local host.
```

```
Switch2#show ip route
R       10.10.0.0/24 [120/1] via 192.168.23.2, 00:30:18, FastEthernet 0/24
R       10.10.1.0/24 [120/1] via 192.168.23.2, 00:30:18, FastEthernet 0/24
C       192.168.0.0/24 is directly connected, Loopback 0
C       192.168.0.1/32 is local host.
C       192.168.1.0/24 is directly connected, Loopback 1
C       192.168.1.1/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/24
C       192.168.23.3/32 is local host.
```

从以上路由表来看，Router1 和 Switch1 通过 RIPv2 收到其他网段 RIPv2 路由信息，Switch2 只收到从 Switch1 发送的 RIPv2 路由更新信息，从 Router1 发送的 RIPv2 更新路由信息被限制了。

9.2.2 分发列表

1. 分发列表概述

路由器分发列表通过访问控制列表控制数据流类型，从而能够利用访问控制列表限制路由选择更新。

使用命令 distribute-list 来配置分发列表，有以下三种方式控制路由选择更新。

（1）入口过滤：过滤进入路由器某接口的更新路由信息。

（2）出口过滤：限制从路由器某接口发送的更新路由信息。

（3）重分发过滤：过滤从另一种路由选择协议重分发的更新路由信息。

如图 9.3 所示，显示了一个基于入站接口和出站接口应用分发列表的路由选择更新过滤过程，具体如下。

（1）路由器接收到或准备发送关于一个或多个网络的路由更新。

（2）路由器查询该操作中所涉及的接口（路由更新进入的接口，或需要通告的路由选择更新的出站接口）。

（3）路由器确定是否有与接口相关联的分发列表。

（4）如果该接口不存在相关联的分发列表，则按正常方式处理分组。

（5）如果该接口有相关联的分发列表，则路由器将扫描分发列表所引用的访问控制列表，以查找

与路由选择更新匹配的条目。

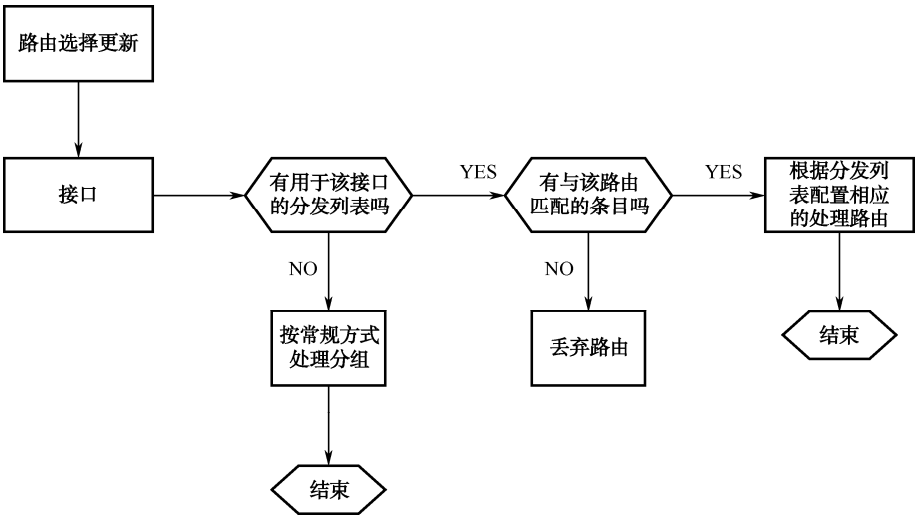


图 9.3　使用分发列表的工作过程

（6）如果访问控制列表中存在匹配的条目，则按配置的方式处理，根据匹配的访问控制列表语句，permit 或 deny 该路由。

（7）如果在访问控制列表中未找到匹配条目，则访问控制列表最后隐含的 deny 策略将导致丢弃该路由。

2. 配置分发列表

配置分发列表分为以下几个步骤。

第 1 步：定义访问控制列表，设定要过滤的网络地址条件。

第 2 步：确定数据流是在入站接口上过滤还是在出站接口过滤，或者是从另一种路由选择协议重分发而来的路由更新。

第 3 步：在 RIP 路由进程配置模式下应用 distribute-list 命令过滤路由信息。

（1）distribute-list（RIP）命令。

由于 RIP 路由协议将路由信息放置在更新报文中进行发送，可以使用 distribute-list 命令对更新路由进行过滤。如果将分发列表应用于路由器接口的输入（in）方向上，可禁止的符合限制条件的路由信息从该接口输入；如果将分发列表应用于路由器接口的输出（out）方向上，将禁止符合限制条件的路由信息从该接口发出，如图 9.4 所示。

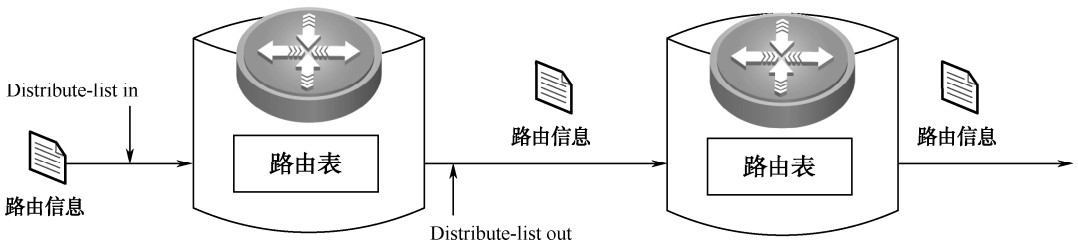


图 9.4　RIP 路由协议应用分发列表

在 RIP 路由协议下，使用 distribute-list 命令对路由信息进行过滤，该命令格式如下：

访问列表编号或名称　　接口类型及编号

ruijie(config-router)#distribute-list *{access-list-number | name}* **{in|out}** *[interface-type interface-number]*

设备名称为ruijie
处于路由配置模式　　数据流方向

其中，*access-list-number* 为指定的访问列表编号，只有访问列表允许的路由，才可以被接收或发送；in 表示对输入方向的更新报文进行过滤；out 表示对输出方式的更新报文进行过滤；*interface-type* 为指定接口的类型，如 Fastethernet，Serial 等；*interface-number* 为指定接口的编号，如 0/0 等。

默认情况下，没有定义任何分发列表。该命令的 no 形式删除已定义的分发列表，如 ruijie(config-router)# no distribute-list 10 in fastethernet 0/0。

提示：*如果没指定接口，就对所有接口接收的路由更新报文进行处理。*

例如：RIP 对从 Fastethernet 0/0 接口接收的路由进行了控制处理，只允许接收 172.16 开头的路由。

```
ruijie(config)#router rip
ruijie(config-router)#network 200.168.23.0
ruijie(config-router)#distribute-list 10 in fastethernet 0/0
ruijie(config-router)#no auto-summary
ruijie(config)#access-list 10 permit 172.16.0.0 0.0.255.255
```

例如：RIP 路由进程只对外通告 192.168.12.0/24 路由。

```
ruijie(config)#router rip
ruijie(config-router)#network 200.4.4.0
ruijie(config-router)#network 192.168.12.0
ruijie(config-router)#distribute-list 10 out
ruijie(config-router)#version 2
ruijie(config)#access-list 10 permit 192.168.12.0 0.0.0.255
```

例如：如图 9.5 所示，两台路由器及一台三层交换机都启动动态路由协议 RIPv2，为了过滤 172.16.0.0 网段，在路由器 Router2 的接口 Fa0/0 入站方向应用分发列表。

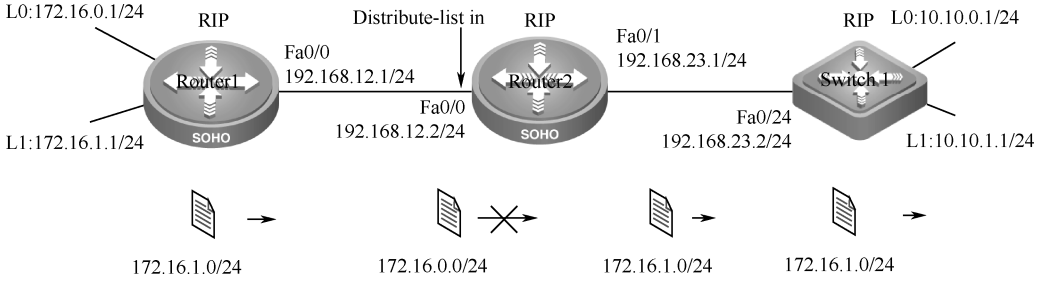


图 9.5　RIP 路由协议应用分发列表实例

设备参数配置如下。

路由器 Router1 参数配置：

```
ruijie(config)#hostname Router1
Router1(config)#interface FastEthernet 0/0
Router1(config-if)#ip address 192.168.12.1 255.255.255.0
Router1(config-if)#no shutdown
Router1(config-if)#exit
```

```
Router1(config)#interface Loopback 0
Router1(config-if)#ip address 172.16.0.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#interface Loopback 1
Router1(config-if)#ip address 172.16.1.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#router rip
Router1(config-router)#version 2
Router1(config-router)#network 172.16.0.0
Router1(config-router)#network 192.168.12.0
Router1(config-router)#no auto-summary
Router1(config-router)#timers basic 5 10 15
Router1(config-router)#end
Router1#
```

路由器 Router2 参数配置：

```
ruijie(config)#hostname Router2
Router2(config)#access-list 10 deny 172.16.0.0 0.0.0.255
Router2(config)#access-list 10 permit any
Router2(config)#interface FastEthernet 0/0
Router2(config-if)#ip address 192.168.12.2 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#interface FastEthernet 0/1
Router2(config-if)#ip address 192.168.23.1 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#router rip
Router2(config-router)#version 2
Router2(config-router)#network 192.168.12.0
Router2(config-router)#network 192.168.23.0
Router2(config-router)#no auto-summary
Router2(config-router)#distribute-list 10 in FastEthernet 0/0
Router2(config-router)#timers basic 5 10 15
Router2(config-router)#end
Router2#
```

交换机 Switch1 参数配置：

```
ruijie(config)#hostname Switch1
Switch1(config)#interface FastEthernet 0/24
Switch1(config-if)#no switchport
Switch1(config-if)#ip address 192.168.23.2 255.255.255.0
Switch1(config-if)#no shutdown
Switch1(config-if)#exit
Switch1(config)#interface Loopback 0
```

```
Switch1(config-if)#ip address 10.10.0.1 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#interface Loopback 1
Switch1(config-if)#ip address 10.10.1.1 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#router rip
Switch1(config-router)#version 2
Switch1(config-router)#network 10.10.0.0
Switch1(config-router)#network 192.168.23.0
Switch1(config-router)#no auto-summary
Switch1(config-router)#timers basic 5 10 15
Switch1(config-router)#end
Switch1#
```

显示路由表：

```
Router1#show ip route
R     10.10.0.0/24 is via 192.168.12.2, 00:00:02, FastEthernet 0/0
R     10.10.1.0/24 is via 192.168.12.2, 00:00:02, FastEthernet 0/0
R     192.168.23.0/24 is via 192.168.12.2, 00:00:02, FastEthernet 0/0
C     172.16.0.0/24 is directly connected, Loopback 0
C     172.16.0.1/32 is local host.
C     172.16.1.0/24 is directly connected, Loopback 1
C     172.16.1.1/32 is local host.
C     192.168.12.0/24 is directly connected, FastEthernet 0/0
C     192.168.12.1/32 is local host.
```

```
Router2#show ip route
C     10.10.0.0/24 is directly connected, Loopback 0
C     10.10.0.1/32 is local host.
C     10.10.1.0/24 is directly connected, Loopback 1
C     10.10.1.1/32 is local host.
R     172.16.1.0/24 [120/1] via 192.168.12.1, 00:00:03, FastEthernet 0/0
R     10.10.0.0/24 is via 192.168.23.2, 00:00:03, FastEthernet 0/1
R     10.10.1.0/24 is via 192.168.23.2, 00:00:03, FastEthernet 0/1
C     192.168.12.0/24 is directly connected, FastEthernet 0/0
C     192.168.12.2/32 is local host.
```

从以上路由表可以看出，由于路由器 Router2 的 Fa0/0 接口 in 方向应用了分发列表，过滤掉了 172.16.0.0 网段路由，三层交换机 Switch1 同样也不会出现该条路由信息。另外，如果在路由器 Router1 的 Fa0/0 接口的 out 方向应用分发列表，也会得到同样效果。

（2）distribute-list（OSPF）命令。

在 OSPF 协议使用命令时，路由更新信息将被放置到链路状态数据库中，而不是放置到路由表中。使用 in 只能过滤本地路由器从链路状态数据库进入到路由表中的路由信息。如果 out，过滤从另一种路由选择协议重分发更新路由信息，也即是本地始发的 5 类 LSA 通告，如图 9.6 所示。

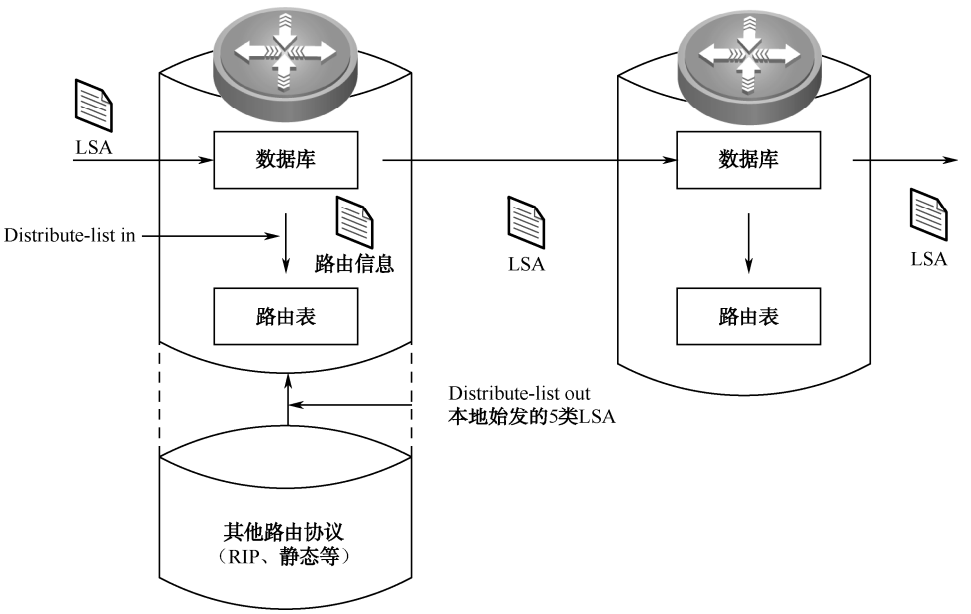


图 9.6　OSPF 路由协议应用分发列表

distribute-list 命令格式如下：

访问列表编号或名称　　接口类型及编号

ruijie(config-router)#distribute-list *{access-list-number | name}* **{in|out}** *[interface-type interface-number]*

设备名称为ruijie，处于路由配置模式　　数据流方向

其中，in 表示对收到的放入链路状态数据库中的 LSA 进行过滤，只有满足过滤条件的才参与 SPF 计算，并计算相应的路由。它并不会影响链路状态数据库，对邻居的路由表也不会产生任何影响，只会影响到本地 OSPF 计算的路由表项。

out 表示对从其他协议重分布到 OSPF 中的路由进行过滤，但它本身不执行路由重分发，一般与 redistribute 命令配合使用。

例如：允许从 Fa0/0 接口接收的 172.16.0.0/16 网络路由参与 SPF 计算，并放入到路由表中。

```
ruijie(config)#access-list 3 permit 172.16.0.0 0.0.255.255
ruijie(config)#router ospf 25
ruijie(config-router)#distribute-list 3 in fastethernet 0/0
```

例如：允许 100.100.0.0/16 网络路由从静态路由重分发到 OSPF 区域。

```
ruijie(config)#access-list 22 permit 100.100.0.0 0.0.255.255
ruijie(config)#router ospf 1
ruijie(config)#redistribute static subnets
ruijie(config-router)#distribute-list 22 out static
```

例如：如图 9.7 所示，两台路由器及一台三层交换机都启动动态路由协议 OSPF，为了过滤 172.16.0.0 网段，在路由器 Router2 的接口 Fa0/0 入站方向应用分发列表。

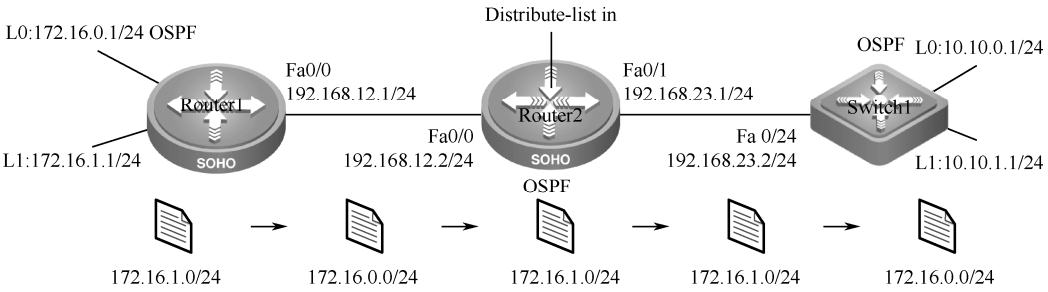


图 9.7　OSPF 路由协议应用分发列表实例

设备参数配置如下。

① 路由器 Router1 参数配置：

```
ruijie(config)#hostname Router1
Router1(config)#interface FastEthernet 0/0
Router1(config-if)#ip address 192.168.12.1 255.255.255.0
Router1(config-if)#no shutdown
Router1(config-if)#exit
Router1(config)#interface Loopback 0
Router1(config-if)#ip address 172.16.0.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#interface Loopback 1
Router1(config-if)#ip address 172.16.1.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#router ospf 1
Router1(config-router)#network 172.16.0.0 0.0.0.255 area 10
Router1(config-router)#network 172.16.1.0 0.0.0.255 area 10
Router1(config-router)#network 192.168.12.0 0.0.0.255 area 0
Router1(config-router)#end
Router1#
```

② 路由器 Router2 参数配置：

```
ruijie(config)#hostname Router2
Router2(config)#access-list 10 deny 172.16.0.0 0.0.0.255
Router2(config)#access-list 10 permit any
Router2(config)#interface FastEthernet 0/0
Router2(config-if)#ip address 192.168.12.2 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#interface FastEthernet 0/1
Router2(config-if)# ip address 192.168.23.1 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#router ospf 1
Router2(config-router)#network 192.168.12.0 0.0.0.255 area 0
Router2(config-router)#network 192.168.23.0 0.0.0.255 area 20
```

```
Router2(config-router)#distribute-list 10 in FastEthernet 0/0
Router2(config-router)#end
Router2#
```

③ 交换机 Switch1 参数配置：

```
ruijie(config)#hostname Switch1
Switch1(config)#interface FastEthernet 0/24
Switch1(config-if)#no switchport
Switch1(config-if)# ip address 192.168.23.2 255.255.255.0
Switch1(config-if)#no shutdown
Switch1(config-if)#exit
Switch1(config)#interface Loopback 0
Switch1(config-if)#ip address 10.10.0.1 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#interface Loopback 1
Switch1(config-if)#ip address 10.10.1.1 255.255.255.0
Switch1(config-if)#exit
Switch1(config)#router ospf 1
Switch1(config-router)# network 10.10.0.0 0.0.0.255 area 20
Switch1(config-router)# network 10.10.1.0 0.0.0.255 area 20
Switch1(config-router)# network 192.168.23.0 0.0.0.255 area 20
Switch1(config-router)#end
Switch1#
```

④ 显示路由表：

```
Router1#show    ip route
O IA 10.10.0.0/24 [110/1] via 192.168.12.2, 00:06:38, FastEthernet 0/0
O IA 10.10.1.0/24 [110/1] via 192.168.12.2, 00:06:38, FastEthernet 0/0
C       172.16.0.0/24 is directly connected, Loopback 0
C       172.16.0.1/32 is local host.
C       172.16.1.0/24 is directly connected, Loopback 0
C       172.16.1.1/32 is local host.
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
O IA 192.168.23.0/24 [110/2] via 192.168.12.2, 00:06:38, FastEthernet 0/0
```

```
Router2#show    ip route
O       10.10.0.1/32 [110/1] via 192.168.23.2, 00:07:15, FastEthernet 0/1
O       10.10.1.1/32 [110/1] via 192.168.23.2, 00:07:15, FastEthernet 0/1
O IA 172.16.1.0/24 [110/1] via 192.168.12.1, 00:06:06, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
```

```
Switch1# show ip route
C       10.10.0.0/24 is directly connected, Loopback 0
C       10.10.0.1/32 is local host.
C       10.10.1.0/24 is directly connected, Loopback 1
C       10.10.1.1/32 is local host.
O IA 172.16.0.0/24 [110/2] via 192.168.23.1, 00:06:37, FastEthernet 0/24
O IA 172.16.1.0/24 [110/2] via 192.168.23.1, 00:06:37, FastEthernet 0/24
C       192.168.23.0/24 is directly connected, FastEthernet 0/24
C       192.168.23.2/32 is local host.
O IA 192.168.12.0/24 [110/2] via 192.168.23.1, 00:06:37, FastEthernet 0/24
```

从以上路由表可以看出，由于路由器 Router2 的 Fa0/0 接口 in 方向应用了分发列表，过滤掉了本路由器的从数据库到路由表的 172.16.0.0 网段路由，三层交换机 Switch1 路由表中仍然有 172.16.0.0 网段路由。原因是 OSPF 协议是链路状态路由协议，所有设备的链路状态数据库内容相同，分发列表只是过滤本地路由表，对其他路由器没有影响。

9.3 扩展知识

1. 管理距离（AD）概述

管理距离（Administrative Distance，AD）是确定一个路由协议可信度的一种度量值。在各种路由协议中，根据可信度不同赋予不同 AD 值，通常 AD 值越小，可信度越高。不同路由协议的默认管理距离值如表 9.1 所示。

表 9.1 默认管理距离值

序号	路 由 源	默认管理距离	说 明
1	Connected interface	0	直接连接接口
2	Static route out an interface	0	静态路由（本地接口输出）
3	Static route to a next hop	1	静态路由（下一跳）
4	External BGP	20	
5	OSPF	110	
6	IS-IS	115	
7	RIP v1, v2	120	
8	Internal BGP	200	
9	Unknown	255	不可达

从表 9.1 可以看出，直接连接网段的路由可信度最高；其次是静态路由；动态路由中 OSPF 协议为 110，RIP 协议为 120，如果一台同时运行 RIP 和 OSPF 协议的路由器，分别从 RIP 和 OSPF 协议获得到达某网段的路由信息，则优先选择从 OSPF 路由协议获得的路由。路由器路由表中的路由都是最佳路由。

如图 9.8 所示，路由器 Router2 分别从动态路由协议 RIP 及 OSPF 获得 172.16.1.0/24 网段路由信息，由于 OSPF 管理距离为 110，优于 RIP 的 120，因此优先选择 OSPF 路由，即从 Router4→Router3→Router2 路径到达 172.16.1.0/24 网络。

如果希望改变路由选择路径，可以改变路由管理距离值。在如图 9.8 所示网络中，将动态路由协议 RIP 的管理距离值设置为 90，低于 OSPF 协议默认管理距离值 110，那么，Router2 将选择从 Router4→Router1→Router2 路径到达 172.16.1.0/24 网络。

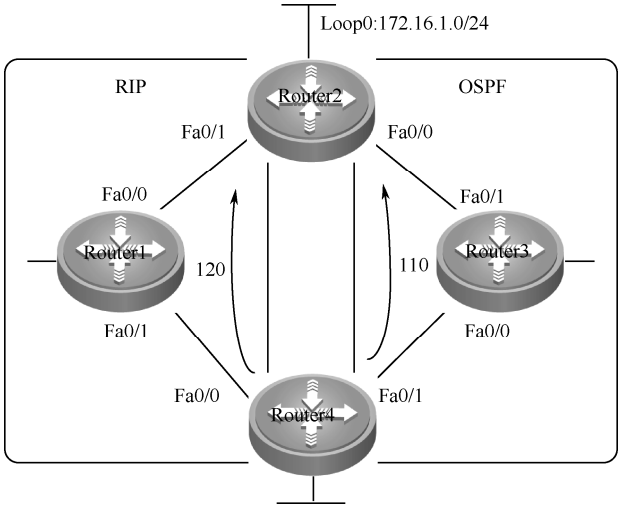


图 9.8　管理距离比较

2. 修改管理距离值

当有多个路由进程通告同一目标网络的路由，特别是在做路由的冗余备份时经常会遇到这种情况，此时可以通过手工修改指定路由的管理距离，达到线路备份的目的。另外通过灵活设定管理距离值，也可以达到路由过滤的效果。

修改 RIP 及 OSPF 协议的管理距离值方法略有不同，现分别叙述如下。

（1）修改 RIP 路由协议的管理距离值。要设置 RIP 路由的管理距离，在 RIP 路由进程模式下使用 distance 命令。该命令格式如下：

设置的管理距离　　*比较通配符*

ruijie(config-router)#distance *distance [ip-address wildcard]*

设备名称为ruijie
处于路由配置模式　　*路由来源IP地址*

其中，*distance* 为设置 RIP 路由管理距离，允许设置<1-255>之间的整数；*ip-address* 为路由来源 IP 地址；*wildcard* 为定义 IP 地址比较通配符，0 表示精确匹配，1 表示不做比较。

默认情况下，RIP 路由协议管理距离值为 120。可以使用该命令的 no 形式恢复默认设置，如 ruijie(config-router)#no distance。

例如：将配置 RIP 路由管理距离为 160，并指定从 192.168.12.1 学习到的路由的管理距离为 123。

```
ruijie(config)#router rip
ruijie(config-router)#distance 160
ruijie(config-router)#distance 123 192.168.12.1 0.0.0.0
```

（2）修改 OSPF 路由协议管理距离值。修改 OSPF 路由协议的管理距离值与修改 RIP 管理距离值不同，OSPF 拥有不同类型的路由，如区域内路由、区域间路由及外部路由等。我们可以设置不同类型路由具有不同的管理距离值。OSPF 路由协议修改管理距离值的命令格式如下：

设置区域间管理距离值

ruijie(config-router)#distance ospf {intra-area *<1-255>* **| inter-area** *<1-255>* **| external***<1-255>***}**

设备名称为ruijie
处于路由配置模式

设置区域内管理距离值

设置外部路由管理距离值

其中，intra-area *<1-255>*为设置区域内路由管理距离，默认值为 110；inter-area *<1-255>*为设置区域间路由管理距离，默认值为 110；external *<1-255>*为设置外部路由管理距离，默认值为 110。使用该命令的 no 选项恢复默认值，如 ruijie(config-router)#no distance ospf。

例如：配置 OSPF 外部路由管理距离为 160。

```
ruijie(config)#router ospf 1
ruijie(config-router)#distance ospf external 160
```

例如：如图 9.9 所示，三台路由器 RA、RB、RC 运行 RIP 路由协议，RA 分别与 RB、RC 连接。通过配置 RA 的 RIP 管理距离，使得 RA 只能学到 RB 通告的路由，并忽略所有 RC 的通告，并且从 RB 学到路由的管理距离为 99。

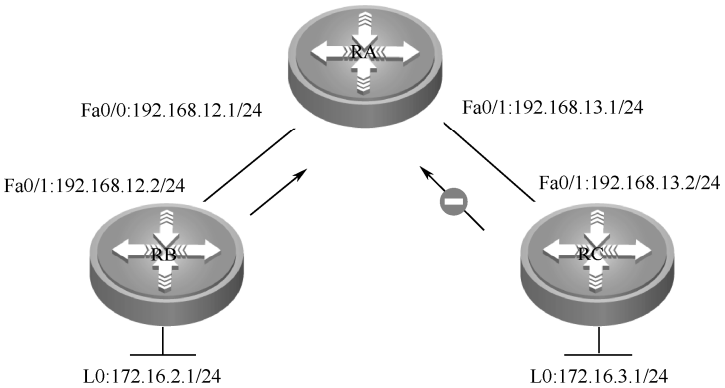


图 9.9　配置 RIP 管理距离值

将路由器 RA 配置为将从 RB 学习的 RIP 路由管理距离值为 99，从其他路由器学到路由的管理距离值全部为 255。

路由器 RA 配置代码如下：

```
RA(config)#interface FastEthernet 0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface FastEthernet 0/1
RA(config-if)#ip address 192.168.13.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#router rip
RA(config-router)#version 2
RA(config-router)#network 192.168.12.0
RA(config-router)#network 192.168.13.0
RA(config-router)#no auto-summary
RA(config-router)#distance 255
```

```
RA(config-router)#distance 99 192.168.12.2 0.0.0.0
RA(config-if)#end
RA#
```

路由器 RB 配置代码如下：

```
RB(config)#interface FastEthernet 0/1
RB(config-if)#ip address 192.168.12.2 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
RB(config)#interface Loopback 0
RB(config-if)#ip address 172.16.2.1 255.255.255.0
RB(config-if)#exit
RB(config)#router rip
RB(config-router)#version 2
RB(config-router)#network 192.168.12.0
RB(config-router)#network 172.16.2.0
RB(config-router)#no auto-summary
RB(config-router)#end
RB#
```

路由器 RC 配置代码如下：

```
RC(config)#interface FastEthernet 0/1
RC(config-if)#ip address 192.168.13.2 255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
RC(config)#interface Loopback 0
RC(config-if)#ip address 172.16.3.1 255.255.255.0
RC(config-if)#exit
RC(config)#router rip
RC(config-router)#version 2
RC(config-router)#network 192.168.13.0
RC(config-router)#network 172.16.3.0
RC(config-router)#no auto-summary
RC(config-router)#end
RC#
```

显示路由器 RA 的路由表如下：

```
RA#show ip route
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
C       192.168.13.0/24 is directly connected, FastEthernet 0/1
C       192.168.13.1/32 is local host.
R       172.16.2.0/24 [99/1] via 192.168.12.2, 00:00:22, FastEthernet 0/0
```

显示路由器 RA 的动态路由协议 RIP 如下：

```
RA#show ip rip
Routing Protocol is "rip"
```

```
Sending updates every 30 seconds, next due in 0 seconds
Invalid after 180 seconds, flushed after 120 seconds
Outgoing update filter list for all interface is: not set
Incoming update filter list for all interface is: not set
Default redistribution metric is 1
Redistributing:
Default version control: send version 2, receive version 2
  Interface              Send  Recv  Key-chain
  FastEthernet 0/0       2     2
  FastEthernet 0/1       2     2
Routing for Networks:
  192.168.12.0
  192.168.13.0
Distance: (default is 255)
```

9.4 演示实例

1. 背景描述

在如图 9.10 所示网络中，三层交换机 Fa0/1 接口以路由口方式连接计算机，VLAN10 中的接口 Fa0/23、Fa0/24 分别连接路由器 R2 和 R3。全部设备启用 RIPv2 路由协议。VLAN10 中的 SVI 接口 IP 地址为 192.168.12.1/24，其他接口 IP 地址如图 9.10 所示。

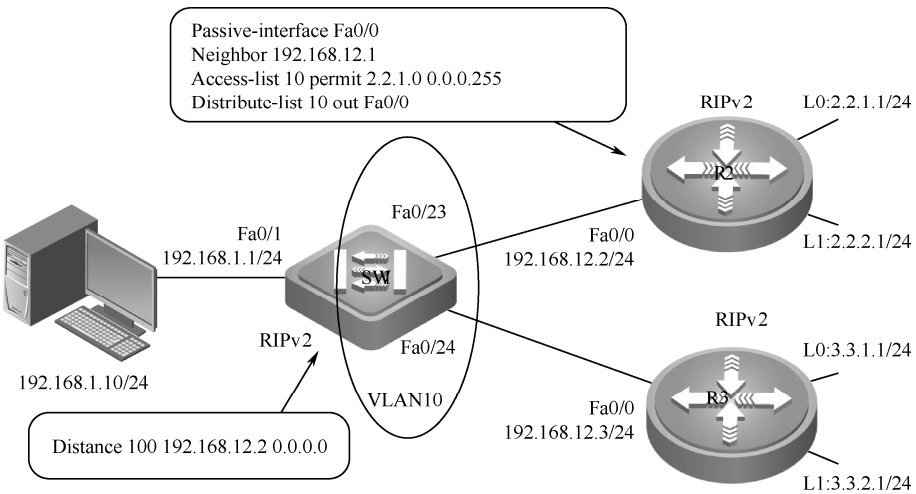


图 9.10　路由选择控制与过滤实例

要求通过适当配置实现如下功能。

（1）路由器 R2 只允许向交换机 SW1 发送路由更新报文。

（2）在路由器 R2 向交换机 SW1 发送的更新报文中只允许包含 2.2.1.0/24 网络路由信息，其他网络路由信息不允许发送。

（3）交换机 SW1 从路由器 R2 收到的 RIP 路由管理距离值设置为 100，从其他设备收到的 RIP 路由管理距离值为默认值 120。

2. 操作步骤

（1）设备参数配置如下。

交换机 SW1 配置代码如下：

```
ruijie(config)#hostname sw1
SW1(config)#VLAN 10
SW1(config-VLAN)#exit
SW1(config)#interface FastEthernet 0/1
SW1(config-if)#no switchport
SW1(config-if)#ip address 192.168.1.1 255.255.255.0
SW1(config-if)#exit
SW1(config)#interface range FastEthernet 0/23-24
SW1(config-range-if)#switchport access VLAN 10
SW1(config-range-if)#exit
SW1(config)#interface VLAN 10
SW1(config-if)#ip address 192.168.10.1 255.255.255.0
SW1(config-if)#exit
SW1(config)#router rip
SW1(config-router)#version 2
SW1(config-router)#network 192.168.1.0
SW1(config-router)#network 192.168.10.0
SW1(config-router)#no auto-summary
SW1(config-router)#distance 100 192.168.12.2 0.0.0.0
SW1(config-router)#end
SW1#
```

路由器 R2 配置代码如下：

```
ruijie(config)#hostname R2
R2(config)#access-list 10 permit 2.2.1.0 0.0.0.255
R2(config)#interface FastEthernet 0/0
R2(config-if)#ip address 192.168.10.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface Loopback 0
R2(config-if)#ip address 2.2.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#interface Loopback 1
R2(config-if)#ip address 2.2.2.1 255.255.255.0
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#passive-interface FastEthernet 0/0
R2(config-router)#network 2.2.1.0 0.0.0.255
R2(config-router)#network 2.2.2.0 0.0.0.255
R2(config-router)#network 192.168.10.0
R2(config-router)#neighbor 192.168.10.1
R2(config-router)#no auto-summary
```

```
R2(config-router)#distribute-list 10 out FastEthernet 0/0
R2(config-router)#end
R2#
```

路由器 R3 配置代码如下：

```
ruijie(config)#hostname R3
R3(config)#interface FastEthernet 0/0
R3(config-if)#ip address 192.168.10.3 255.255.255.0
R3(config-if)#exit
R3(config)#interface Loopback 0
R3(config-if)#ip address 3.3.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#interface Loopback 1
R3(config-if)#ip address 3.3.2.1 255.255.255.0
R3(config-if)#exit
R3(config)#router rip
R3(config-router)#version 2
R3(config-router)#network 3.3.1.0 0.0.0.255
R3(config-router)#network 3.3.2.0 0.0.0.255
R3(config-router)#network 192.168.10.0
R3(config-router)#no auto-summary
R3(config-router)#end
R3#
```

（2）效果验证。

查看交换机划分 VLAN 情况如下：

```
SW1#show VLAN
VLAN Name                               Status    Ports
---------------                         ----------  -------------------------------------
1 VLAN0001                              STATIC    Fa0/2, Fa0/3, Fa0/4, Fa0/5
                                                  Fa0/6, Fa0/7, Fa0/8, Fa0/9
                                                  Fa0/10, Fa0/11, Fa0/12, Fa0/13
                                                  Fa0/14, Fa0/15, Fa0/16, Fa0/17
                                                  Fa0/18, Fa0/19, Fa0/20, Fa0/21
                                                  Fa0/22, Gi0/25, Gi0/26, Gi0/27
                                                  Gi0/28
10 VLAN0010                             STATIC    Fa0/23, Fa0/24
SW1#
```

从上述显示结果看出，Fa0/23、Fa0/24 接口包含在 VLAN10 中。

查看交换机路由表如下：

```
SW1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
R       2.2.1.0/24 [100/1] via 192.168.10.2, 00:14:53, VLAN 10
R       3.3.1.0/24 [120/1] via 192.168.10.3, 00:29:21, VLAN 10
R       3.3.2.0/24 [120/1] via 192.168.10.3, 00:29:21, VLAN 10
C       192.168.1.0/24 is directly connected, FastEthernet 0/1
C       192.168.1.1/32 is local host.
C       192.168.10.0/24 is directly connected, VLAN 10
C       192.168.10.1/32 is local host.
SW1#
```

从上述显示结果看出，交换机只学到 R2 路由器的 2.2.1.0/24 网络的路由信息。

查看路由器 R2 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C       2.2.1.0/24 is directly connected, Loopback 0
C       2.2.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
R       3.3.1.0/24 [120/1] via 192.168.10.3, 00:28:07, FastEthernet 0/0
R       3.3.2.0/24 [120/1] via 192.168.10.3, 00:28:07, FastEthernet 0/0
R       192.168.1.0/24 [120/1] via 192.168.10.1, 00:20:30, FastEthernet 0/0
C       192.168.10.0/24 is directly connected, FastEthernet 0/0
C       192.168.10.2/32 is local host.
```

从以上显示结果看出，路由器 R2 能够学习到 R3 及 SW1 设备的路由信息。

查看路由器 R3 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
```

```
C       3.3.1.0/24 is directly connected, Loopback 0
C       3.3.1.1/32 is local host.
C       3.3.2.0/24 is directly connected, Loopback 1
C       3.3.2.1/32 is local host.
R       192.168.1.0/24 [120/1] via 192.168.10.1, 00:19:08, FastEthernet 0/0
C       192.168.10.0/24 is directly connected, FastEthernet 0/0
C       192.168.10.3/32 is local host.
```

从以上显示结果看出，路由器 R3 只能学习到 SW1 的路由信息，不能学习到 R2 的路由信息。

9.5 训练任务

1. 背景描述

如图 9.11 所示，网络中设备分别启动 RIP 及 OSPF 路由协议，为了禁止 Router1 中的 172.16.0.0/24 网段路由发布到其他设备，在 ABR 路由器 Router2 上设置分发列表，限制 172.16.0.0 路由在 OSPF 域内扩散。

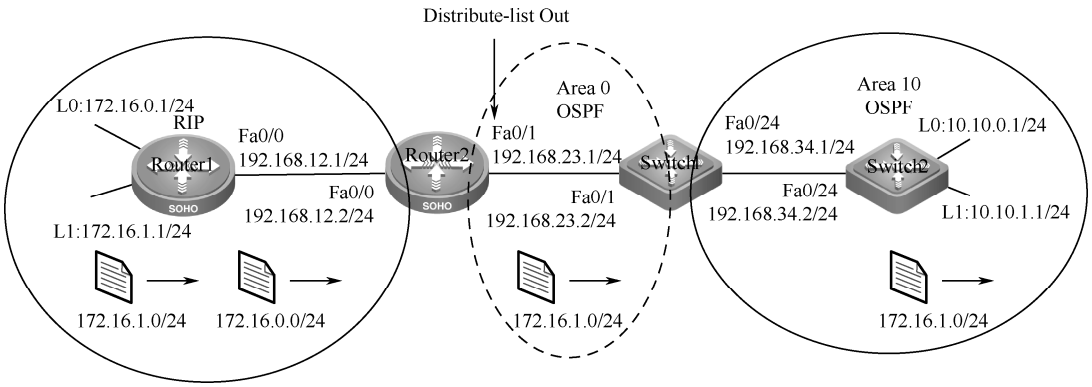


图 9.11　路由重发布应用分发列表

2. 操作提示

设备参数配置如下。

路由器 Router1 配置代码如下：

```
ruijie(config)#hostname Router1
Router1(config)#interface FastEthernet 0/0
Router1(config-if)#ip address 192.168.12.1 255.255.255.0
Router1(config-if)#no shutdown
Router1(config-if)#exit
Router1(config)#interface Loopback 0
Router1(config-if)#ip address 172.16.0.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#interface Loopback 1
Router1(config-if)#ip address 172.16.1.1 255.255.255.0
Router1(config-if)#exit
Router1(config)#router rip
```

```
Router1(config-router)#version 2
Router1(config-router)#network 172.16.0.0
Router1(config-router)#network 192.168.12.0
Router1(config-router)#no auto-summary
Router1(config-router)#timers basic 5 10 15
Router1(config-router)#end
Router1#
```

路由器 Router2 配置代码如下：

```
ruijie(config)#hostname Router2
Router2(config)#access-list 10 deny 172.16.0.0 0.0.0.255
Router2(config)#access-list 10 permit any
Router2(config)#interface FastEthernet 0/0
Router2(config-if)#ip address 192.168.12.2 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#interface FastEthernet 0/1
Router2(config-if)#ip address 192.168.23.1 255.255.255.0
Router2(config-if)#no shutdown
Router2(config-if)#exit
Router2(config)#router rip
Router2(config-router)#version 2
Router2(config-router)#network 192.168.12.0
Router2(config-router)#redistribute connected
Router2(config-router)#redistribute ospf
Router2(config-router)#no auto-summary
Router2(config-router)#timers basic 5 10 15
Router2(config-router)#end
Router2(config)#router ospf 1
Router2(config-router)#redistribute connected
Router2(config-router)#redistribute rip subnets
Router2(config-router)#network 192.168.23.0 0.0.0.255 area 0
Router2(config-router)#distribute-list 10 out rip
Router2(config-router)#end
Router2#
```

交换机 Switch1 代码配置如下：

```
ruijie(config)#hostname Switch1
Switch1(config)#interface FastEthernet 0/1
Switch1(config-if)#no switchport
Switch1(config-if)#ip address 192.168.23.2 255.255.255.0
Switch1(config-if)#no shutdown
Switch1(config-if)#exit
Switch1(config)#interface FastEthernet 0/24
Switch1(config-if)#no switchport
```

```
Switch1(config-if)#ip address 192.168.34.1 255.255.255.0
Switch1(config-if)#no shutdown
Switch1(config-if)#exit
Switch1(config)#router ospf 1
Switch1(config-router)#network 192.168.23.0 0.0.0.255 area 0
Switch1(config-router)#network 192.168.34.0 0.0.0.255 area 10
Switch1(config-router)#end
Switch1#
```

交换机 Switch2 代码配置如下：

```
ruijie(config)#hostname Switch2
Switch2(config)#interface FastEthernet 0/24
Switch2(config-if)#no switchport
Switch2(config-if)#ip address 192.168.34.2 255.255.255.0
Switch2(config-if)#no shutdown
Switch2(config-if)#exit
Switch2(config)#interface Loopback 0
Switch2(config-if)#ip address 10.10.0.1 255.255.255.0
Switch2(config-if)#no shutdown
Switch2(config-if)#exit
Switch2(config)#interface Loopback 1
Switch2(config-if)#ip address 10.10.1.1 255.255.255.0
Switch2(config-if)#no shutdown
Switch2(config-if)#exit
Switch2(config)#router ospf 1
Switch2(config-router)#network 10.10.0.0 0.0.0.255 area 10
Switch2(config-router)#network 10.10.1.0 0.0.0.255 area 10
Switch2(config-router)#network 192.168.34.0 0.0.0.255 area 10
Switch2(config-router)#end
Switch2#
```

显示路由表：

```
Router1#show ip route
R       10.10.0.1/32 [120/1] via 192.168.12.2, 00:00:03, FastEthernet 0/0
R       10.10.1.1/32 [120/1] via 192.168.12.2, 00:00:03, FastEthernet 0/0
C       172.16.0.0/24 is directly connected, Loopback 0
C       172.16.0.1/32 is local host.
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
R       192.168.23.0/24 [120/1] via 192.168.12.2, 00:00:03, FastEthernet 0/0
R       192.168.34.0/24 [120/1] via 192.168.12.2, 00:00:03, FastEthernet 0/0
```

```
Router2#show ip route
```

```
O IA 10.10.0.1/32 [110/2] via 192.168.23.2, 00:38:00, FastEthernet 0/1
O IA 10.10.1.1/32 [110/2] via 192.168.23.2, 00:37:45, FastEthernet 0/1
R    172.16.0.0/24 [120/1] via 192.168.12.1, 00:00:03, FastEthernet 0/0
R    172.16.1.0/24 [120/1] via 192.168.12.1, 00:00:03, FastEthernet 0/0
C    192.168.12.0/24 is directly connected, FastEthernet 0/0
C    192.168.12.2/32 is local host.
C    192.168.23.0/24 is directly connected, FastEthernet 0/1
C    192.168.23.1/32 is local host.
O IA 192.168.34.0/24 [110/2] via 192.168.23.2, 00:40:25, FastEthernet 0/1
```

```
switch1#show ip route
O    10.10.0.1/32 [110/1] via 192.168.34.2, 00:38:31, FastEthernet 0/24
O    10.10.1.1/32 [110/1] via 192.168.34.2, 00:38:16, FastEthernet 0/24
O E2 172.16.1.0/24 [110/20] via 192.168.23.1, 00:29:42, FastEthernet 0/1
O E2 192.168.12.0/24 [110/20] via 192.168.23.1, 00:30:33, FastEthernet 0/1
C    192.168.23.0/24 is directly connected, FastEthernet 0/1
C    192.168.23.2/32 is local host.
C    192.168.34.0/24 is directly connected, FastEthernet 0/24
C    192.168.34.1/32 is local host.
```

```
switch2#show ip route
C    10.10.0.0/24 is directly connected, Loopback 0
C    10.10.0.1/32 is local host.
C    10.10.1.0/24 is directly connected, Loopback 1
C    10.10.1.1/32 is local host.
O E2 172.16.1.0/24 [110/20] via 192.168.34.1, 00:30:14, FastEthernet 0/24
O E2 192.168.12.0/24 [110/20] via 192.168.34.1, 00:31:05, FastEthernet 0/24
O IA 192.168.23.0/24 [110/2] via 192.168.34.1, 00:39:21, FastEthernet 0/24
C    192.168.34.0/24 is directly connected, FastEthernet 0/24
C    192.168.34.2/32 is local host.
```

从以上路由表情况来看，由于在路由器 Router2 上配置了分发列表，并使用 distribute-list out 命令应用于 OSPF 路由协议中，过滤了本地网络 5 类 LSA 通告，使得 Switch1 和 Switch2 的链路状态数据库中没有 172.16.0.0/24 网络路由信息，自然路由表中就没有此条信息了。

如果 Switch1 和 Switch2 上应用分发列表 distribute-list out 命令，由于不是“本地”，不能过滤 5 类 LSA 通告，链路状态数据库仍然存在 172.16.0.0/24 网络路由信息。应用分发列表 distribute-list in 命令只是过滤本地路由表，对其他路由器没有影响。

此处的“本地”是指进行路由重分发的路由器。

练习题

1. 选择题

（1）下面关于被动接口的描述正确的是（　　）。

A．被动接口只发生更新报文，不接收更新报文
B．被动接口只接收更新报文，不发送更新报文
C．被动接口既接收又发送更新报文
D．被动接口既不接收也不发送更新报文

（2）下面设置被动接口命令正确的是（　　）。
A．ruijie#passive-interface Fa 0/1
B．ruijie(config)#passive-interface Fa 0/1
C．ruijie(config-router)#passive-interface Fa 0/1
D．ruijie(config-if)#passive-interface Fa 0/1

（3）分发列表命令 distribute-list 10 in fastethernet 0/0 表示的含义是（　　）。
A．在接口 Fastethernet 0/0 中设置的序号为 10 的分发列表
B．在接口 Fastethernet 0/0 中设置共计 10 分发列表
C．通过 ACL 10 定义限制条件，应用在接口 Fastethernet 0/0 输入方向的分发列表
D．通过 ACL 10 定义限制条件，应用除接口 Fastethernet 0/0 之外所有接口的分发列表

（4）分发列表过滤的是（　　）类型数据。
A．用户发送的应用报文
B．系统发送的系统报文
C．路由协议发送的路由更新报文
D．设备发送的设备故障报文

（5）下面修改 AD 值命令正确的是（　　）。
A．ruijie(config-router)#distance 100 192.168.12.1 0.0.0.0
B．ruijie(config)#distance 100
C．ruijie#distance 100 192.168.12.1 0.0.0.0
D．ruijie(config-if)#distance 100 192.168.12.1 0.0.0.0

2. 简答题

（1）分别描述被动接口及分发列表主要功能。
（2）描述分发列表与 ACL 的区别。

项目 10　路由重发布

网络工程中，由于公司购并或其他原因导致使用的网络设备动态路由协议不止一种，如同时配置 RIP 及 OSPF 协议。为了网络路由资源共享，不同路由协议间需要路由迁移，相关路由协议间交换路由信息。路由重发布技术因此得到广泛应用。

本章主要介绍路由重发布的相关知识及配置路由重发布的方法。

知识点、技能点

1. 了解路由重发布的定义及应用;
2. 掌握路由重发布的配置技能。

10.1　问题提出

路由器支持多种路由协议，每种路由协议都有自己的特性，如 RIP 度量值采用跳数、支持有类路由；OSPF 度量值采用 Cost、支持无类路由。如何实现 RIP 与 OSPF 协议间路由迁移，使得在不同路由域间学习到对方的路由，只有采用路由重发布技术。

10.2　相关知识

1. 路由重发布概述

（1）路由重发布的定义。路由重发布是在连接到不同路由域的边界路由器上实现不同路由域（如 RIP 或 OSPF）之间交换和通告路由选择信息的能力。

（2）路由重发布的应用环境。网络中很多场合需要采用路由重发布技术，主要有以下几种情况。

- 因公司购并造成公司网络整合，且两家公司网络应用不同路由协议。如果要统一使用一种路由协议，重新配置工作量很大，且需要网络停止工作，实际上是不允许的。
- 因公司某种业务或策略需要，配置多种路由协议。
- 有些厂商设备只支持 RIP 或 OSPF 协议，多厂商设备使用也需要配置路由重发布。

（3）路由重发布的条件。路由重发布的前提是两个路由协议必须在一个网络体系结构内，如 TCP/IP。如果一个在 TCP/IP 内，另一个在 IPX 内，二者是不允许路由重发布的。

（4）路由重发布的过程。如果未配置路由重发布功能，路由协议发布路由更新时，在路由域内只通告通过其进程获得的路由信息，路由进程间不共享路由信息，如图 10.1 所示。通常将这种路由器路由进程间不共享路由信息的方式称为夜航式路由选择。

夜航式路由选择方式各路由协议不能学习到其他路由域内的路由。在如图 10.1 所示网络中，RIP 路由协议不能学习到 172.16.1.0～172.16.3.0 网络路由，OSPF 也不能学习到 10.1.1.0～10.1.3.0 网络的路由。

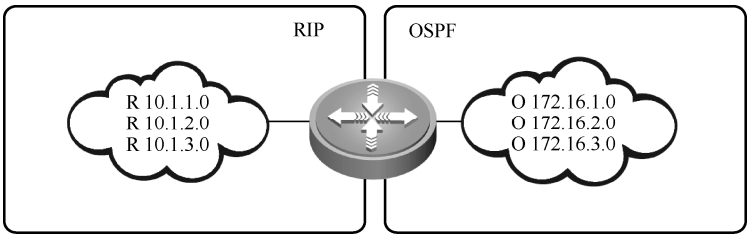


图 10.1　夜航式路由选择

如果在自治系统边界路由器（ASBR）上配置路由重发布，使得路由协议可以学习到其他路由域的路由信息，通过这种方式学习到的路由作为外部路由，路由器优先选择内部路由，如图 10.2 所示。

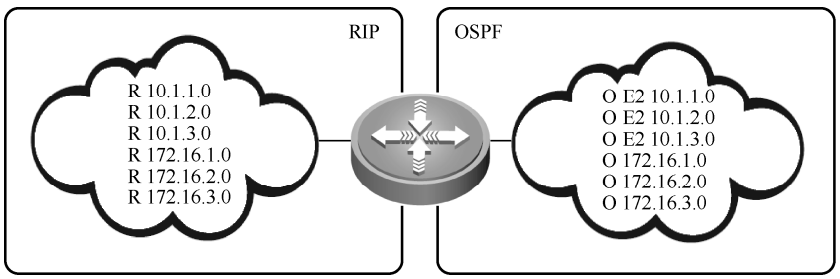


图 10.2　路由重发布过程

2. 配置路由重发布

路由重发布时，将一个路由协议的量度转换成另一种路由协议的量度是没有必要的，因为不同路由协议所采用的量度计算方法是完全不同的。RIP 量度计算是基于跳数的，而 OSPF 是基于带宽的，因此它们计算出来的量度是没有可比性的。但是路由重发布时，又必须要设置一个象征性的量度，否则路由重发布将失败。

（1）重发布到 RIP 路由域的路由。重发布到 RIP 路由就是在 RIP 路由协议进程中，接收其他外部路由信息，包括 OSPF 路由、直连路由、静态路由等。默认情况下，重发布到 RIP 路由域的度量值为 1，也可以通过命令修改默认度量值。

① 修改 RIP 路由进程的默认度量值。如果重发布一个路由协议进程时，没有明确定义度量值，RIP 将使用 default-metric 定义的度量值；如果有明确定义度量值，该值将覆盖 default-metric 定义的度量值。

修改默认度量值的命令格式如下：

ruijie(config-router)#default-metric *metric*

设备名称为ruijie
处于路由配置模式

设置的默认度量值

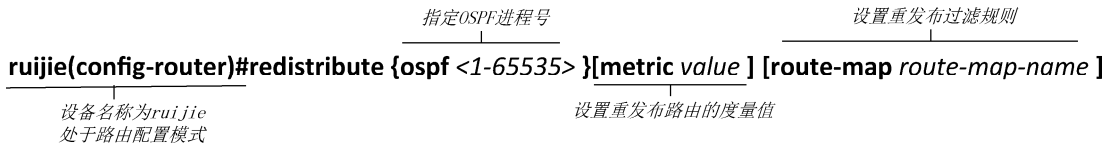
其中，*metric* 为设置的默认度量值。默认情况下，*metric* 为 1。使用该命令的 no 形式恢复默认值，如 ruijie(config-router)#no default-metric。

例如：将 RIP 路由协议对从 OSPF 路由协议学到的路由进行了重新发布，这些重发布路由的初始 RIP 度量值设置为 3。

```
ruijie(config)#router rip
ruijie(config-router)#default-metric 3
```

```
ruijie(config-router)#redistribute ospf 100
```

② 将 OSPF 路由域中的路由重发布到 RIP 路由域中，是经常使用的路由重发布方式。该命令格式如下：



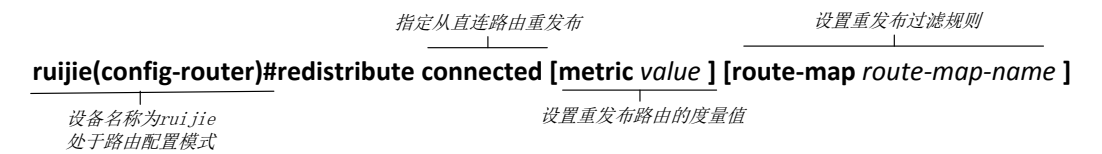
其中，*<1-65535>*为指定的 OSPF 进程号；*value* 为重发布路由的度量值，该值将优先于使用 default-metric 命令定义的默认度量值；*route-map-name* 为设置重发布路由过滤规则的路由图名称，如在 route-map-name 中使用 set 命令修改了度量值，该度量值将优先于使用 metric 命令设置的优先级。总之，各设置度量值的优先级为 route-map-name>metric value>default-metric。

要取消已设置的重发布外部路由，使用该命令的 no 形式，如 ruijie(config-router)#no redistribute ospf 1 metric 2。

例如：将 OSPF 路由进程 1 的路由发布到 RIP 中，且度量值为 2 跳。

```
ruijie(config)#router rip
ruijie(config-router)#redistribute ospf 1 metric 2
```

③ 将直连路由重发布到 RIP 路由域中的命令格式如下：

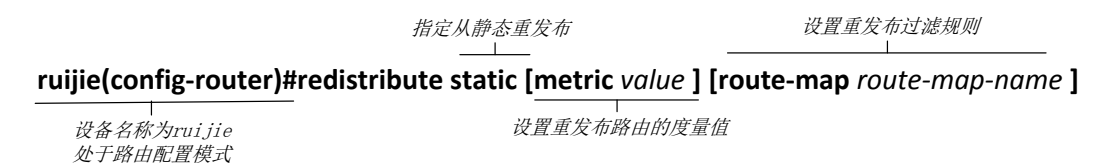


要取消已设置的重发布外部路由，使用该命令的 no 形式，如 ruijie(config-router)#no redistribute connected metric 2。

例如：将直连路由发布到 RIP 中，且度量值为 2 跳。

```
ruijie(config)#router rip
ruijie(config-router)#redistribute connected metric 2
```

④ 将静态路由重发布到 RIP 路由域中的命令格式如下：

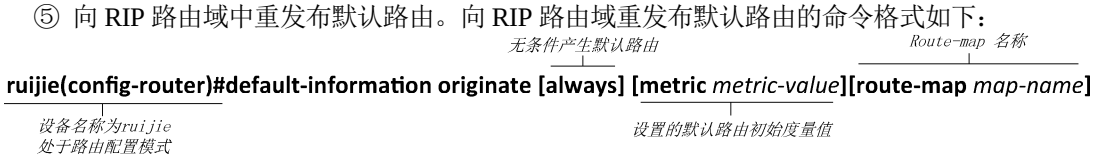


要取消已设置的重发布外部路由，使用该命令的 no 形式，如 ruijie(config-router)#no redistribute static metric 2。

例如：将静态路由发布到 RIP 中，且度量值为 2 跳。

```
ruijie(config)#router rip
ruijie(config-router)#redistribute static metric 2
```

⑤ 向 RIP 路由域中重发布默认路由。向 RIP 路由域重发布默认路由的命令格式如下：



其中，**always**（可选）为该关键字使得 RIP 无条件产生默认路由，不管本地是否存在默认路由；**metric** *metric-value*（可选）为默认路由初始的度量值，metric-value 取值范围为 1～15，默认情况下，度量值为 1；**route-map** *map-name*（可选）为关联的 Route-map 的名字，默认没有关联 route-map。

使用该命令的 no 形式取消已生成的默认路由，如 ruijie(config-router)#no default-information originate always。

例如：产生一条默认路由到 RIP 路由域。

```
ruijie(config)#router rip
ruijie(config-router)#default-information originate always
```

（2）重发布到 OSPF 路由域的路由。重发布到 OSPF 路由就是在 OSPF 路由协议进程中，接收其他外部路由信息，包括 RIP 路由、直连路由、静态路由等。默认情况下，重发布到 OSPF 路由域的度量值为 20，也可以通过命令修改默认度量值。

① 修改 RIP 路由进程的默认度量值。

如果重发布一个路由协议进程时，没有明确定义度量值，OSPF 将使用 default-metric 定义的度量值；如果有明确定义度量值，该值将覆盖 default-metric 定义的度量值。

修改默认度量值的命令格式如下：

ruijie(config-router)#default-metric *metric*

设备名称为ruijie，处于路由配置模式　　*设置的默认度量值*

其中，*metric* 为设置的默认度量值。默认情况下，*metric* 为 20。使用该命令的 no 形式恢复默认值，如 ruijie(config-router)#no default-metric。

例如：将 OSPF 重发布路由的初始化度量值设为 50。

```
ruijie(config)#router rip
ruijie(config-router)#network 192.168.12.0
ruijie(config-router)#version 2
ruijie(config-router)#exit
ruijie(config)#router ospf
ruijie(config-router)#network 172.16.10.0 0.0.0.255 area 0
ruijie(config-router)#default-metric 50
ruijie(config-router)#redistribute rip subnets
```

② 将 RIP 路由域中的路由重发布到 OSPF 路由域中，也是经常使用的路由重发布方式。该命令格式如下：

指定RIP　　*设置重发布过滤规则的路由图名称*

ruijie(config-router)#redistribute RIP [metric *value* **] [route-map** *route-map-name***] [subnets]**

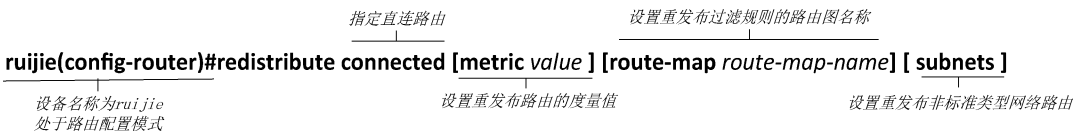
设备名称为ruijie，处于路由配置模式　　*设置重发布路由的度量值*　　*设置重发布非标准类型网络路由*

其中，*value* 为重发布路由的度量值，该值将优先于使用 default-metric 命令定义的默认度量值；*route-map-name* 为设置重发布路由过滤规则的路由图名称；subnets 为可重发布非标准类网络，因为 OSFP 是无类路由协议，而 RIP 是有类的路由协议，所以在重发布的时候需要指定关键字 subnets。

例如：将 RIP 路由发布到 OSPF 中，且度量值为 100，使用 subnets 选项允许发布子网路由。

```
ruijie(config)#router ospf 1
ruijie(config-router)#redistribute rip metric 100 subnets
```

③ 将直连路由重发布到 OSPF 路由域中的命令格式如下：

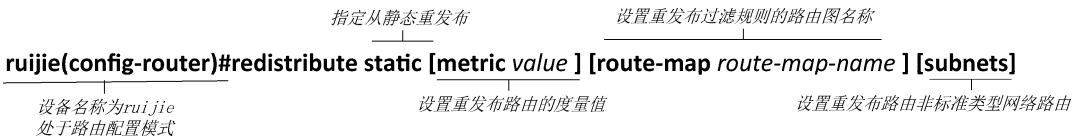


要取消已设置的重发布外部路由，使用该命令的 no 形式，如 ruijie(config-router)#no redistribute connected metric 30 subnets。

例如：将直连路由发布到 OSPF 中，且度量值为 50。

```
ruijie(config)#router ospf 1
ruijie(config-router)#redistribute connected metric 50 subnets
```

④ 将静态路由重发布到 OSPF 路由域中的命令格式如下：

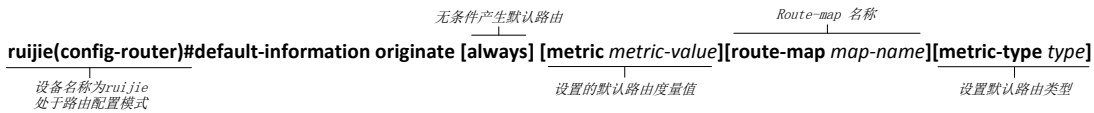


要取消已设置的重发布外部路由，使用该命令的 no 形式，如 ruijie(config-router)#no redistribute static metric 30。

例如：将直连路由发布到 OSPF 中，且度量值为 30 跳。

```
ruijie(config)#router ospf 1
ruijie(config-router)#redistribute static metric 30 subnets
```

⑤ 向 OSPF 路由域中重发布默认路由。向 OSPF 路由域重发布默认路由的命令格式如下：



其中，**always**（可选）为该关键字使得 RIP 无条件产生默认路由，不管本地是否存在默认路由。

metric *metric-value*（可选）为默认路由初始的度量值，metric-value 取值范围为 1～15，默认情况下，度量值为 1。

route-map *map-name*（可选）为关联的 Route-map 的名字，默认没有关联 Route-map。

使用该命令的 no 形式取消已生成的默认路由，如 ruijie(config-router)#no default-information originate always。

metric-type *type*（可选）为默认路由的类型。OSPF 外部路由有两种类型：类型 1，不同路由设备上看到的度量值不一样；类型 2，所有路由设备看到的度量值都一样。类型 1 的外部路由比类型 2 可信。默认为类型 2。

例如：产生一条默认路由到 OSPF 路由域。

```
ruijie(config)#router ospf 1
ruijie(config-router)#default-information originate always
```

3. 应用实例

在如图 10.3 所示网络中，在路由器 R2 上配置路由重发布，实现将直连路由、静态路由、RIP 动态路由重发布到 OSPF 路由域，将直连路由、静态路由、OSPF 路由重发布到 RIP 路由域。

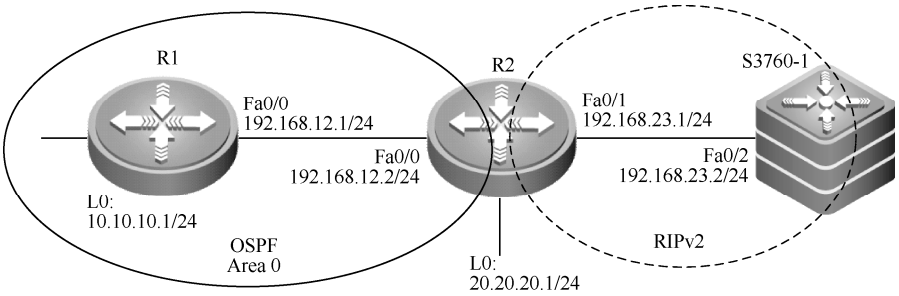


图 10.3　路由重发布应用实例

（1）设备参数配置步骤。

路由器 R1 配置如下：

```
ruijie(config)#hostname R1
R1(config)#interface FastEthernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface Loopback 0
R1(config-if)#ip address 10.10.10.1 255.255.255.0
R1(config-if)#exit
R1(config)#router ospf 1
R1(config-router)#network 10.10.10.0 0.0.0.255 area 0
R1(config-router)#network 192.168.12.0 0.0.0.255 area 0
R1(config-router)#end
R1#
```

路由器 R2 配置如下：

```
ruijie(config)#hostname R2
R2(config)#interface FastEthernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface FastEthernet 0/1
R2(config-if)#ip address 192.168.23.1 255.255.255.0
R2(config-if)#exit
R2(config)#interface Loopback 0
R2(config-if)#ip address 20.20.20.1 255.255.255.0
R2(config-if)#exit
R2(config)#router ospf 1
R2(config-router)#redistribute connected subnets
R2(config-router)#redistribute static subnets
R2(config-router)#redistribute rip
R2(config-router)#network 192.168.12.0 0.0.0.255 area 0
R2(config-router)#exit
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#network 192.168.23.0
```

```
R2(config-router)#no auto-summary
R2(config-router)#redistribute connected
R2(config-router)#redistribute static
R2(config-router)#redistribute ospf 1
R2(config-router)#exit
R2(config)#ip route 100.100.1.0 255.255.255.0 192.168.23.2
R2(config)#ip route 200.200.1.0 255.255.255.0 192.168.12.1
R2(config)#end
R2#
```

交换机 S3760-1 配置如下：

```
ruijie(config)#hostname S3760-1
S3760-1(config)#interface FastEthernet 0/24
S3760-1(config-if)#no switchport
S3760-1(config-if)#ip address 192.168.23.2 255.255.255.0
S3760-1(config-if)#exit
S3760-1(config)#router rip
S3760-1(config-router)#version 2
S3760-1(config-router)#network 192.168.23.0
S3760-1(config-router)#no auto-summary
S3760-1(config-router)#end
S3760-1#
```

（2）结果测试。

显示路由器 R1 的路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       10.10.10.0/24 is directly connected, Loopback 0
C       10.10.10.1/32 is local host.
O E2 20.20.20.0/24 [110/20] via 192.168.12.2, 00:05:20, FastEthernet 0/0
O E2 100.100.1.0/24 [110/20] via 192.168.12.2, 00:05:02, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
O E2 192.168.23.0/24 [110/20] via 192.168.12.2, 00:07:03, FastEthernet 0/0
```

从上述显示结果可以看出，路由器 R1 从 R2 学到直连路由 20.20.20.0/24、静态路由 100.100.1.0/24 及 RIP 路由 192.168.23.0/24 等，度量值都是默认值 20。

显示路由器 R2 的路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
```

```
            O - OSPF, IA - OSPF inter area
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
O     10.10.10.1/32 [110/1] via 192.168.12.1, 00:33:39, FastEthernet 0/0
C     20.20.20.0/24 is directly connected, Loopback 0
C     20.20.20.1/32 is local host.
S     100.100.1.0/24 [1/0] via 192.168.23.2
C     192.168.12.0/24 is directly connected, FastEthernet 0/0
C     192.168.12.2/32 is local host.
C     192.168.23.0/24 is directly connected, FastEthernet 0/1
C     192.168.23.1/32 is local host.
S     200.200.1.0/24 [1/0] via 192.168.12.1
```

从上述显示结果可以看出，路由器 R2 从 R1 学到 OSPF 路由 10.10.10.1/32 路由。

显示三层交换机 S3760-1 的路由表如下：

```
S3760-1#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
            O - OSPF, IA - OSPF inter area
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
R     10.10.10.1/32 [120/1] via 192.168.23.1, 00:21:04, FastEthernet 0/24
R     20.20.20.0/24 [120/1] via 192.168.23.1, 00:04:57, FastEthernet 0/24
R     192.168.12.0/24 [120/1] via 192.168.23.1, 00:21:04, FastEthernet 0/24
C     192.168.23.0/24 is directly connected, FastEthernet 0/24
C     192.168.23.2/32 is local host.
R     200.200.1.0/24 [120/1] via 192.168.23.1, 00:04:57, FastEthernet 0/24
```

从上述显示结果可以看出，三层交换机 S3760-1 从路由器 R2 学到直连路由 20.20.20.0/24、静态路由 200.200.1.0/24 及 OSPF 路由 10.10.10.1/32 等，度量值为默认值 1。

提问：路由器 R1 为什么没有学习到静态路由 200.200.1.0/24 呢？路由器 R2 为什么没有学习到静态路由 100.100.1.0/24 呢？

10.3 扩展知识

次优路由问题

在进行路由重发布配置时，有时会出现次优路由问题，下面通过一个实例进行具体描述。

某企业网络中同时运行 RIPv2 和 OSPF 两种路由协议，两个路由域通过两台边界路由 RB 和 RC 互连，网络拓扑如图 10.4 所示。为了使两个路由域能够共享路由信息，在 RB 上进行 RIPv2 和 OSPF 的双向重分发，RC 作为内部网络出口，向 RIP 和 OSPF 中生成并通告了一条默认路由。

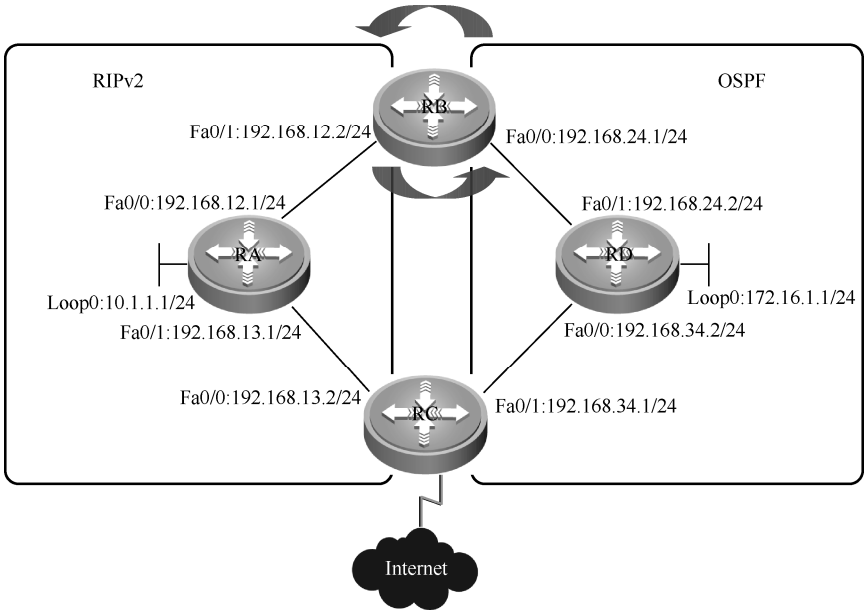


图 10.4　某企业网络拓扑图

在如图 10.4 所示网络中，通过路由重发布实现路由，但是在路由器 RB 进行双向的重分发后产生了一个路由选择问题。路由器 RC 同时通过 RIP 和 OSPF 获得了到达 RIP 网络 10.1.1.0/24 的路由，下一跳分别是路由器 RA 和 RD。但是由于 OSPF 的 AD 值小于 RIP 的 AD 值，所有 RC 将优选 OSPF 路由，即通过 RD 到达 RIP 网络，显然产生了次优路径选择问题，如图 10.5 所示。您作为网管应如何解决此问题呢。

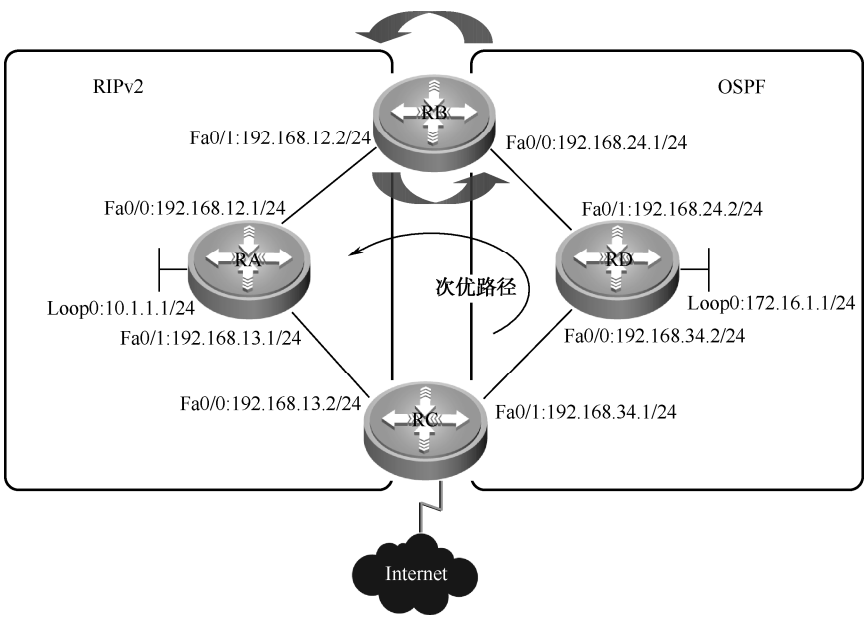


图 10.5　网络产生次优路径

由于 OSPF 协议的 AD 值小于 RIP 的 AD 值，所以路由器 RC 会选择通过 RD 的路径。在这种情

况下要避免次优路径的选择，可以在 RC 上调整 OSPF 协议外部路由的 AD 值，使其大于 RIP 的 AD 值，这样 RC 将选择通过 RA 去往 RIP 网络 10.1.1.0/24，如图 10.6 所示。

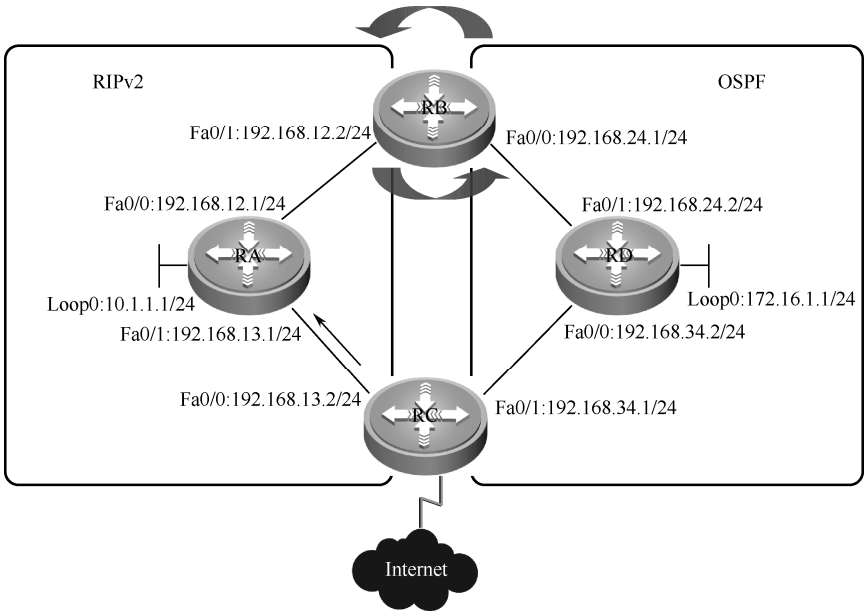


图 10.6　调整 AD 后网络路由

设备配置代码如下。

第 1 步：基本配置。

```
RA(config)#interface FastEthernet 0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface FastEthernet 0/1
RA(config-if)#ip address 192.168.13.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface Loop0
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
```

```
RB(config)#interface FastEthernet 0/0
RB(config-if)#ip address 192.168.24.1 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
RB(config)#interface FastEthernet 0/1
RB(config-if)#ip address 192.168.12.2 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
```

```
RC(config)#interface FastEthernet 0/0
RC(config-if)#ip address 192.168.13.2 255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
RC(config)#interface FastEthernet 0/1
RC(config-if)#ip address 192.168. 34.1 255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
```

```
RD(config)#interface FastEthernet 0/0
RD(config-if)#ip address 192.168.34.2 255.255.255.0
RD(config-if)#no shutdown
RD(config-if)#exit
RD(config)#interface FastEthernet 0/1
RD(config-if)#ip address 192.168.24.2 255.255.255.0
RD(config-if)#no shutdown
RD(config-if)#exit
RD(config)#interface Loop0
RD(config-if)#ip address 172.16.1.1 255.255.255.0
RD(config-if)#no shutdown
RD(config-if)#exit
```

第 2 步：配置 RIP 及 OSPF 协议。

```
RA(config)#router rip
RA(config-router)#version 2
RA(config-router)#network 10.0.0.0
RA(config-router)#network 192.168.12.0
RA(config-router)#network 192.168.13.0
RA(config-router)#no auto-summary
```

```
RB(config)#router rip
RB(config-router)#version 2
RB(config-router)#network 192.168.12.0
RB(config-router)#no auto-summary
RB(config-router)#exit
RB(config)#router ospf 1
RB(config-router)#network 192.168.24.0 0.0.0.255 area 0
RB(config-router)#end
```

```
RC(config)#router rip
RC(config-router)#version 2
RC(config-router)#network 192.168.13.0
RC(config-router)#default-information originate
RC(config-router)#no auto-summary
```

```
RC(config-router)#exit
RC(config)#router ospf 1
RC(config-router)#network 192.16.34.0 0.0.0.255 area 0
RC(config-router)#default-information originate always
RC(config-router)#end
```

```
RD(config)#router ospf 1
RD(config-router)#network 172.16.1.0 0.0.0.255 area 0
RD(config-router)#network 192.168.24.0 0.0.0.255 area 0
RD(config-router)#network 192.168.34.0 0.0.0.255 area 0
RD(config-router)#end
```

第 3 步：配置路由重发布。

```
RB(config)#router rip
RB(config-router)#redistribute connected
RB(config-router)#redistribute ospf 1
RB(config-router)#default-information originate always
RB(config-router)#exit
RB(config)#router ospf 1
RB(config-router)#redistribute connected subnets
RB(config-router)#redistribute rip subnets
RB(config-router)#default-information originate
RB(config-router)#end
```

```
RA#show ip route
Codes:      C - connected, S - static,    R - RIP B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
        * - candidate default
Gateway of last resort is 192.168.12.2 to network 0.0.0.0
R*      0.0.0.0/0 [120/1] via 192.168.12.2, 00:00:11, FastEthernet 0/0
C       10.1.1.0/24 is directly connected, Loopback 0
C       10.1.1.1/32 is local host.
R       172.16.1.1/32 [120/1] via 192.168.12.2, 00:00:11, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
C       192.168.13.0/24 is directly connected, FastEthernet 0/1
C       192.168.13.1/32 is local host.
R       192.168.24.0/24 [120/1] via 192.168.12.2, 00:00:11, FastEthernet 0/0
R       192.168.34.0/24 [120/1] via 192.168.12.2, 00:00:11, FastEthernet 0/0
```

```
RB#show ip route
```

```
Codes:      C - connected, S - static,    R - RIP B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
        * - candidate default
Gateway of last resort is no set
R       10.1.1.0/24 [120/1] via 192.168.12.1, 00:00:12, FastEthernet 0/1
O       172.16.1.1/32 [110/1] via 192.168.24.2, 00:16:58, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.2/32 is local host.
R       192.168.13.0/24 [120/1] via 192.168.12.1, 00:00:12, FastEthernet 0/1
C       192.168.24.0/24 is directly connected, FastEthernet 0/0
C       192.168.24.1/32 is local host.
O       192.168.34.0/24 [110/2] via 192.168.24.2, 00:18:16, FastEthernet 0/0
```

```
RC#show ip route
Codes:      C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.34.2 to network 0.0.0.0
O*E2 0.0.0.0/0 [110/1] via 192.168.34.2, 00:01:27, FastEthernet 0/1
O E2 10.1.1.0/24 [110/20] via 192.168.34.2, 00:13:28, FastEthernet 0/1
O       172.16.1.1/32 [110/1] via 192.168.34.2, 00:13:29, FastEthernet 0/1
O E2 192.168.12.0/24 [110/20] via 192.168.34.2, 00:02:16, FastEthernet 0/1
C       192.168.13.0/24 is directly connected, FastEthernet 0/0
C       192.168.13.2/32 is local host.
O       192.168.24.0/24 [110/2] via 192.168.34.2, 00:13:29, FastEthernet 0/1
C       192.168.34.0/24 is directly connected, FastEthernet 0/1
C       192.168.34.1/32 is local host.
```

```
RD#show ip route
Codes:      C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.24.1 to network 0.0.0.0
O*E2 0.0.0.0/0 [110/1] via 192.168.24.1, 00:01:55, FastEthernet 0/1
```

```
O E2 10.1.1.0/24 [110/20] via 192.168.24.1, 00:15:31, FastEthernet 0/1
C       172.16.1.0/24 is directly connected, Loopback 0
C       172.16.1.1/32 is local host.
O E2 192.168.12.0/24 [110/20] via 192.168.24.1, 00:02:44, FastEthernet 0/1
O E2 192.168.13.0/24 [110/20] via 192.168.24.1, 00:15:31, FastEthernet 0/1
C       192.168.24.0/24 is directly connected, FastEthernet 0/1
C       192.168.24.2/32 is local host.
C       192.168.34.0/24 is directly connected, FastEthernet 0/0
C       192.168.34.2/32 is local host.
```

从上述路由器 RC 的路由表看出，RC 去往 10.1.1.0/24 网络的路由，要经过路由器 RD 到达，明显是次优路由。为了解决此问题，需要按如下方式调整进入 RC 路由器 OSPF 的 AD 值。

第 4 步：修改 AD 值。

修改 RC 路由器 OSPF 协议的 AD 值为 140，大于 RIP 的默认值 120。

```
RC(config)#router ospf 1
RC(config-router)#distance ospf external 140
RC(config-router)#end
```

第 5 步：测试。

通过 show ip route 命令，观察在 AD 值修改前后路由器 RC 的路由表变化情况。

```
RC#show ip route
Codes:      C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.13.1 to network 0.0.0.0
R*      0.0.0.0/0 [120/1] via 192.168.13.1, 00:00:58, FastEthernet 0/0
R       10.1.1.0/24 [120/1] via 192.168.13.1, 00:00:58, FastEthernet 0/0
O       172.16.1.1/32 [110/1] via 192.168.34.2, 00:02:07, FastEthernet 0/1
R       192.168.12.0/24 [120/1] via 192.168.13.1, 00:00:58, FastEthernet 0/0
C       192.168.13.0/24 is directly connected, FastEthernet 0/0
C       192.168.13.2/32 is local host.
O       192.168.24.0/24 [110/2] via 192.168.34.2, 00:02:07, FastEthernet 0/1
C       192.168.34.0/24 is directly connected, FastEthernet 0/1
C       192.168.34.1/32 is local host.
```

从上述路由器 RC 的路由表看出，去往 10.1.1.0/24 网络经过路由器 RA 到达。

10.4 演示实例

1. 背景描述

如图 10.7 所示，网络中路由 RIP 及 OSPF 间路由重发布。网络中有三台路由器 RA、RB、RC，其中，RA 和 RB 的接口 Fa0/0 位于 RIPv2 进程中，RC 和 RB 的接口 Fa0/1 位于 OSPF 进程中。

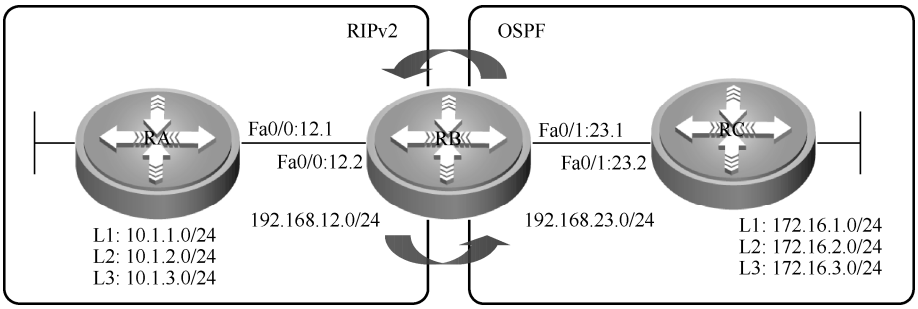


图 10.7　RIP 与 OSPF 间路由重发布实例

RB 需要将 RIP 路由重发布到 OSPF 进程中，并将种子度量值指定为 100，这样 RC 从路由器 RB 中获得前往网络 10.1.1.0、10.1.2.0、10.1.3.0 的三条路由。

RB 也需要将 OSPF 路由发布到 RIP 进程中，并将种子度量值指定为 2，这样 RA 从路由器 RB 中获得前往网络 172.16.1.0、172.16.2.0、172.16.3.0 的三条路由。

2. 操作步骤

设备参数配置如下。

第 1 步：基本配置。

```
RA(config)#interface fastethernet 0/0
RA(config-if)#ip address 192.168.12.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface loopback 1
RA(config-if)#ip address 10.1.1.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loopback 2
RA(config-if)#ip address 10.1.2.1 255.255.255.0
RA(config-if)#exit
RA(config)#interface loopback 3
RA(config-if)#ip address 10.1.3.1 255.255.255.0
RA(config-if)#exit
```

```
RB(config)#interface fastethernet 0/0
RB(config-if)#ip address 192.168.12.2 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
RB(config)#interface fastethernet 0/1
RB(config-if)#ip address 192.168.23.1 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
```

```
RC(config)#interface fastethernet 0/1
RC(config-if)#ip address 192.168.23.2 255.255.255.0
```

```
RC(config-if)#no shutdown
RC(config-if)#exit
RC(config)#interface loopback 1
RC(config-if)#ip address 172.16.1.1 255.255.255.0
RC(config-if)#exit
RC(config)#interface loopback 2
RC(config-if)#ip address 172.16.2.1 255.255.255.0
RC(config-if)#exit
RC(config)#interface loopback 3
RC(config-if)#ip address 172.16.3.1 255.255.255.0
RC(config-if)#exit
```

第 2 步：配置路由协议。

```
RA(config)#router rip
RA(config-router)#version 2
RA(config-router)#network 10.0.0.0
RA(config-router)#network 192.168.12.0
RA(config-router)#no auto-summary
RA(config-router)#exit
```

```
RB(config)#router rip
RB(config-router)#version 2
RB(config-router)#network 192.168.12.0
RB(config-router)#no auto-summary
RB(config-router)#exit
RB(config)#router ospf 1
RB(config-router)#network 192.168.23.0 0.0.0.255 area 0
RB(config-router)#exit
RB(config)#ip route 100.1.1.0 255.255.255.0    Fa 0/1
RB(config-router)#exit
```

```
RC(config)#router ospf 1
RC(config-router)#network 192.168.23.0 0.0.0.255 area 0
RC(config-router)#network 172.16.1.0 0.0.0.255 area 0
RC(config-router)#network 172.16.2.0 0.0.0.255 area 0
RC(config-router)#network 172.16.3.0 0.0.0.255 area 0
RC(config-router)#exit
```

第 3 步：配置路由重发布。

```
RB(config)#router rip
RB(config-router)#redistribute ospf 1 metric 2
RB(config-router)#exit
RB(config)#router ospf 1
RB(config-router)#redistribute rip metric 100 subnets
RB(config-router)#exit
```

第 4 步：查看路由表。

```
RA#show ip route
C       10.1.1.0/24 is directly connected, Loopback 1
C       10.1.1.1/32 is local host.
C       10.1.2.0/24 is directly connected, Loopback 2
C       10.1.2.1/32 is local host.
C       10.1.3.0/24 is directly connected, Loopback 3
C       10.1.3.1/32 is local host.
R       172.16.1.1/32 [120/2] via 192.168.12.2, 00:00:19, FastEthernet 0/0
R       172.16.2.1/32 [120/2] via 192.168.12.2, 00:00:19, FastEthernet 0/0
R       172.16.3.1/32 [120/2] via 192.168.12.2, 00:00:19, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
```

```
RC#show ip route
O E2 10.1.1.0/24 [110/100] via 192.168.23.1, 00:02:35, FastEthernet 0/1
O E2 10.1.2.0/24 [110/100] via 192.168.23.1, 00:02:35, FastEthernet 0/1
O E2 10.1.3.0/24 [110/100] via 192.168.23.1, 00:02:35, FastEthernet 0/1
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.2/32 is local host.
```

在上面的实例中，路由器 RA 没有学习到 192.168.23.0/24 网络路由；同样，路由器 RC 也没有学习到 192.168.12.0/24 网络路由。造成此现象的原因是 192.168.12.0/24 通过 Fa0/0 接口与 RB 直接连接；192.168.23.0/24 通过 Fa0/1 接口与 RB 也直接连接。去往这两个网络的路由属于直连路由，需要将直连路由重发布到 RIP 或 OSPF 中。

由于没有配置静态路由重发布，路由器 RA 和 RC 也没有学习到 100.1.1.0/24 网络路由。

为了使路由器 RA 、RC 能够学习到静态及直连路由，需要完成如下补充配置代码。

```
RB(config)#router rip
RB(config-router)#redistribute connected
RB(config-router)#redistribute static
RB(config-router)#exit
RB(config)#router ospf 1
RB(config-router)#redistribute connected subnets
RB(config-router)#redistribute static subnets
RB(config-router)#exit
```

完成上述配置后，再一次查看路由器 RA 及 RC 的路由表如下：

```
RA#show ip route
C       10.1.1.0/24 is directly connected, Loopback 1
```

```
C       10.1.1.1/32 is local host.
C       10.1.2.0/24 is directly connected, Loopback 2
C       10.1.2.1/32 is local host.
C       10.1.3.0/24 is directly connected, Loopback 3
C       10.1.3.1/32 is local host.
R       100.1.1.0/24 [120/1] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.1.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.2.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.3.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
R       192.168.23.0/24 [120/1] via 192.168.12.2, 00:00:25, FastEthernet 0/0
```

可以看出路由器 RA 已经学到直连路由及静态路由了。

```
RC#show ip route
O E2 10.1.1.0/1 [110/100] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 10.1.2.0/1 [110/100] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 10.1.3.0/1 [110/100] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 100.1.1.0/1 [110/20] via 192.168.23.1, 00:01:19, FastEthernet 0/1
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
O E2 192.168.12.0/1 [110/20] via 192.168.23.1, 00:15:37, FastEthernet 0/1
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.2/32 is local host.
```

可以看出路由器 RC 已经学到直连路由及静态路由了。

如果在如图 10.8 所示网络中路由器 RB 作为接入 Internet 设备，还需要配置默认路由，且将默认路由分发到 RIP 及 OSPF 路由域中的其他路由器中。

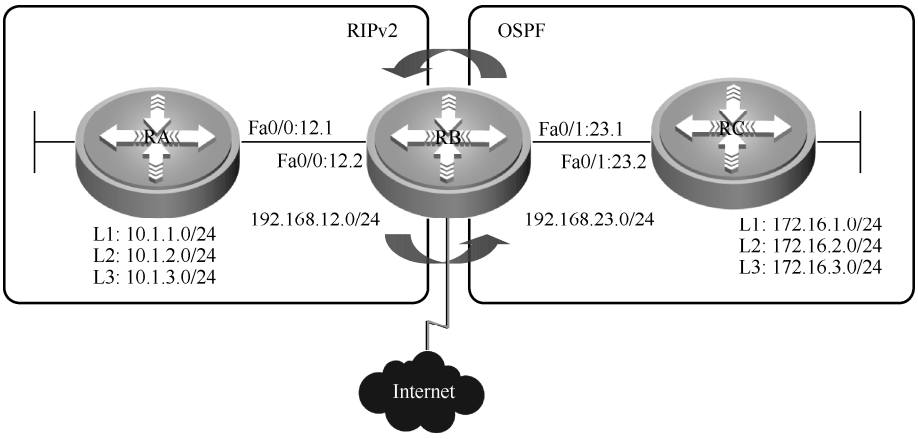


图 10.8　重发布默认路由网络

对于如图 10.8 所示网络中路由器 RB 的默认路由的配置如下：

```
RB(config)#router ospf 1
RB(config-router)#redistribute connected subnets
RB(config-router)#redistribute static subnets
RB(config-router)#redistribute rip subnets
RB(config-router)#network 192.168.23.0 0.0.0.255 area 0
RB(config-router)#default-information originate always
RB(config-router)#exit
RB(config)#router rip
RB(config-router)#version 2
RB(config-router)#network 192.168.12.0
RB(config-router)#no auto-summary
RB(config-router)#redistribute connected
RB(config-router)#redistribute static
RB(config-router)#redistribute ospf
RB(config-router)#default-information originate
RB(config-router)#exit
RB(config)#ip route    100.1.1.0 255.255.255.0    FastEthernet 0/1
```

完成上述配置后，再次查看路由器 RA 及 RC 的路由表。

```
RA#show ip route
Codes:      C - connected, S - static,    R - RIP B - BGP
        O - OSPF, IA - OSPF inter area
                N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
                E1 - OSPF external type 1, E2 - OSPF external type 2
                        i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
                * - candidate default
Gateway of last resort is 192.168.12.2 to network 0.0.0.0
R*      0.0.0.0/0 [120/1] via 192.168.12.2, 00:00:25, FastEthernet 0/0
C       10.1.1.0/24 is directly connected, Loopback 1
C       10.1.1.1/32 is local host.
C       10.1.2.0/24 is directly connected, Loopback 2
C       10.1.2.1/32 is local host.
C       10.1.3.0/24 is directly connected, Loopback 3
C       10.1.3.1/32 is local host.
R       100.1.1.0/24 [120/1] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.1.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.2.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
R       172.16.3.1/32 [120/2] via 192.168.12.2, 00:00:25, FastEthernet 0/0
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.1/32 is local host.
R       192.168.23.0/24 [120/1] via 192.168.12.2, 00:00:25, FastEthernet 0/0
```

可以看出路由器 RA 已经从 RB 中学习到一条默认路由信息了。

```
RC#show ip route
```

```
Codes:    C - connected, S - static, R - RIP, B - BGP
                O - OSPF, IA - OSPF inter area
      N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
                E1 - OSPF external type 1, E2 - OSPF external type 2
           i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
           ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.23.1 to network 0.0.0.0
O*E2 0.0.0.0/0 [110/1] via 192.168.23.1, 00:15:11, FastEthernet 0/1
O E2 10.1.1.0/24 [110/20] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 10.1.2.0/24 [110/20] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 10.1.3.0/24 [110/20] via 192.168.23.1, 00:20:10, FastEthernet 0/1
O E2 100.1.1.0/24 [110/20] via 192.168.23.1, 00:01:19, FastEthernet 0/1
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
O E2 192.168.12.0/24 [110/20] via 192.168.23.1, 00:15:37, FastEthernet 0/1
C       192.168.23.0/24 is directly connected, FastEthernet 0/24
C       192.168.23.2/32 is local host.
```

可以看出路由器 RC 也从 RB 中学习到一条默认路由信息了。

10.5 训练任务

1. 背景描述

在如图 10.9 所示网络中，通过适当配置实现从 RIP 重发布到 OSPF 中路由初始度量值为 30，从 OSPF 重发布到 RIP 中路由初始度量值为 5；三层交换机 SW1 及 SW2 中 Loop1 路由采用直连路由重发布，且初始度量值 SW1 为 50，SW2 为 5；路由器 R3 对于 RIP 及 OSPF 路由域采用默认路由重发布；最终实现内网互通，并能访问 Internet。

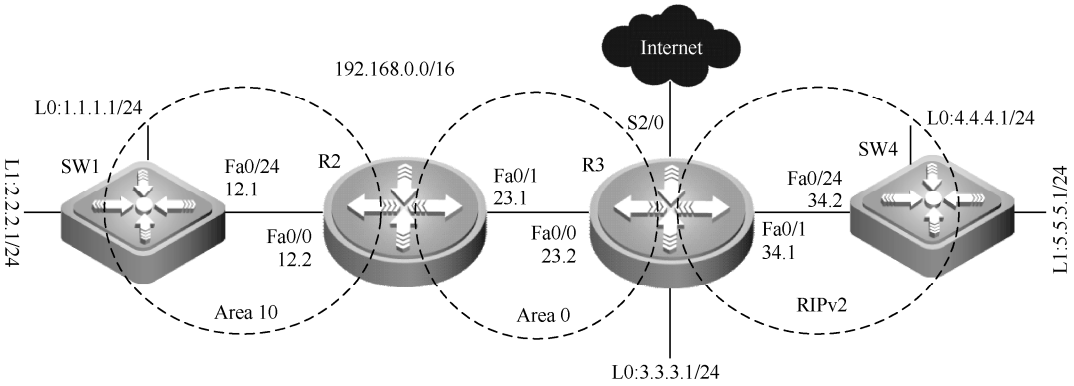


图 10.9　应用案例

2. 操作步骤

（1）设备参数配置。

交换机 SW1 配置如下：

```
ruijie(config)#hostname sw1
SW1(config)#interface FastEthernet 0/24
SW1(config-if)#no switchport
SW1(config-if)#ip address 192.168.12.1 255.255.255.0
SW1(config-if)#exit
SW1(config)#interface Loopback 0
SW1(config-if)#ip address 1.1.1.1 255.255.255.0
SW1(config-if)#exit
SW1(config)#interface Loopback 1
SW1(config-if)#ip address 2.2.2.1 255.255.255.0
SW1(config-if)#exit
SW1(config)#router ospf 1
SW1(config-router)#network 1.1.1.0 0.0.0.255 area 10
SW1(config-router)#network 192.168.12.0 0.0.0.255 area 10
SW1(config-router)#redistribute connected metric 50 subnets
SW1(config-router)#end
SW1#
```

路由器 R2 配置如下：

```
ruijie(config)#hostname R2
R2(config)#interface FastEthernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface FastEthernet 0/1
R2(config-if)#ip address 192.168.23.1 255.255.255.0
R2(config-if)#exit
R2(config)#router ospf 1
R2(config-router)#network 192.168.12.0 0.0.0.255 area 10
R2(config-router)#network 192.168.23.0 0.0.0.255 area 0
R2(config-router)#end
R2#
```

路由器 R3 配置如下：

```
ruijie(config)#hostname R3
R3(config)#interface Serial 2/0
R3(config-if)#ip address 100.1.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#interface FastEthernet 0/0
R3(config-if)#ip address 192.168.23.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface FastEthernet 0/1
R3(config-if)#ip address 192.168.34.1 255.255.255.0
```

```
R3(config-if)#exit
R3(config)#interface Loopback 0
R3(config-if)#ip address 3.3.3.1 255.255.255.0
R3(config-if)#exit
R3(config)#router ospf 1
R3(config-router)#network 192.168.23.0 0.0.0.255 area 0
R3(config-router)#redistribute connected subnets
R3(config-router)#redistribute static subnets
R3(config-router)#redistribute rip metric 30 subnets
R3(config-router)#default-information originate
R3(config-router)#exit
R3(config)#router rip
R3(config-router)#version 2
R3(config-router)#network 192.168.34.0
R3(config-router)#no auto-summary
R3(config-router)#redistribute connected
R3(config-router)#redistribute static
R3(config-router)#redistribute ospf 1 metric 5
R3(config-router)#default-information originate
R3(config-router)#exit
R3(config)#ip route 0.0.0.0 0.0.0.0 100.1.1.2
R3(config)#end
R3#
```

交换机 SW4 配置如下：

```
ruijie(config)#hostname sw4
SW4(config)#interface FastEthernet 0/24
SW4(config-if)#no switchport
SW4(config-if)#ip address 192.168.34.2 255.255.255.0
SW4(config-if)#exit
SW4(config)#interface Loopback 0
SW4(config-if)#ip address 4.4.4.1 255.255.255.0
SW4(config-if)#exit
SW4(config)#interface Loopback 1
SW4(config-if)#ip address 5.5.5.1 255.255.255.0
SW4(config-if)#exit
SW4(config)#router rip
SW4(config-router)#version 2
SW4(config-router)#network 4.4.4.0 0.0.0.255
SW4(config-router)#network 192.168.34.0
SW4(config-router)#no auto-summary
SW4(config-router)#redistribute connected metric 5
SW4(config-router)#end
SW4#
```

（2）结果测试。

显示交换机 SW1 的路由表如下：

```
SW1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.12.2 to network 0.0.0.0
O*E2 0.0.0.0/0 [110/1] via 192.168.12.2, 00:49:45, FastEthernet 0/24
C       1.1.1.0/24 is directly connected, Loopback 0
C       1.1.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
O E2 3.3.3.0/24 [110/20] via 192.168.12.2, 01:14:23, FastEthernet 0/24
O E2 4.4.4.0/24 [110/30] via 192.168.12.2, 00:15:28, FastEthernet 0/24
O E2 5.5.5.0/24 [110/30] via 192.168.12.2, 00:04:01, FastEthernet 0/24
O E2 100.1.1.0/24 [110/20] via 192.168.12.2, 01:14:23, FastEthernet 0/24
C       192.168.12.0/24 is directly connected, FastEthernet 0/24
C       192.168.12.1/32 is local host.
O IA 192.168.23.0/24 [110/2] via 192.168.12.2, 01:24:09, FastEthernet 0/24
O E2 192.168.34.0/24 [110/20] via 192.168.12.2, 01:12:49, FastEthernet 0/24
```

从上述显示结果看出，交换机 SW1 从路由器 R2 中学习到默认路由、RIP 重发布过来的 4.4.4.0/24 和 5.5.5.0/24 路由、直连路由 3.3.3.0/24 和 192.168.34.0/24 及 OSPF 区域间路由 192.168.23.0/24。

显示路由器 R2 的路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.23.2 to network 0.0.0.0
O*E2 0.0.0.0/0 [110/1] via 192.168.23.2, 00:50:59, FastEthernet 0/1
O       1.1.1.0/24 [110/1] via 192.168.12.1, 01:25:21, FastEthernet 0/0
O E2 2.2.2.0/24 [110/50] via 192.168.12.1, 00:03:24, FastEthernet 0/0
O E2 3.3.3.0/24 [110/20] via 192.168.23.2, 01:15:37, FastEthernet 0/1
O E2 4.4.4.0/24 [110/30] via 192.168.23.2, 00:16:42, FastEthernet 0/1
O E2 5.2.5.0/24 [110/30] via 192.168.23.2, 00:12:00, FastEthernet 0/1
O E2 100.1.1.0/24 [110/20] via 192.168.23.2, 01:15:37, FastEthernet 0/1
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
```

```
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
O E2 192.168.34.0/24 [110/20] via 192.168.23.2, 01:14:02, FastEthernet 0/1
```

从上述显示结果看出，路由器 R2 从交换机 SW1 中学习 OSPF 协议 1.1.1.0/24 路由及直连路由 2.2.2.0/24；从路由器 R3 中学习到默认路由、RIP 重发布过来的 4.4.4.0/24 和 5.5.5.0/24 路由、直连路由 192.168.34.0/24。

显示路由器 R3 的路由表如下：

```
R3#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
           O - OSPF, IA - OSPF inter area
           N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
           E1 - OSPF external type 1, E2 - OSPF external type 2
           i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
           ia - IS-IS inter area, * - candidate default
Gateway of last resort is 100.1.1.2 to network 0.0.0.0
S*     0.0.0.0/0 [1/0] via 100.1.1.2
O IA 1.1.1.0/24 [110/2] via 192.168.23.1, 01:26:26, FastEthernet 0/1
O E2 2.2.2.0/24 [110/50] via 192.168.23.1, 00:04:31, FastEthernet 0/1
C       3.3.3.0/24 is directly connected, Loopback 0
C       3.3.3.1/32 is local host.
R       4.4.4.0/24 [120/1] via 192.168.34.2, 00:17:51, FastEthernet 0/0
R       5.5.5.0/24 [120/5] via 192.168.34.2, 00:11:45, FastEthernet 0/0
C       100.1.1.0/24 is directly connected, Serial 2/0
C       100.1.1.1/32 is local host.
O IA 192.168.12.0/24 [110/2] via 192.168.23.1, 01:26:29, FastEthernet 0/1
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.2/32 is local host.
C       192.168.34.0/24 is directly connected, FastEthernet 0/0
C       192.168.34.1/32 is local host.
```

显示交换机 SW2 路由表如下：

```
S3760B#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
           O - OSPF, IA - OSPF inter area
           N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
           E1 - OSPF external type 1, E2 - OSPF external type 2
           i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
           ia - IS-IS inter area, * - candidate default
Gateway of last resort is 192.168.34.1 to network 0.0.0.0
R*     0.0.0.0/0 [120/1] via 192.168.34.1, 00:53:08, FastEthernet 0/24
R       1.1.1.0/24 [120/2] via 192.168.34.1, 00:27:01, FastEthernet 0/24
R       2.2.2.0/24 [120/2] via 192.168.34.1, 00:07:01, FastEthernet 0/24
R       3.3.3.0/24 [120/1] via 192.168.34.1, 00:27:01, FastEthernet 0/24
C       4.4.4.0/24 is directly connected, Loopback 0
```

```
C       4.4.1.1/32 is local host.
C       5.5.5.0/24 is directly connected, Loopback 1
C       5.5.5.1/32 is local host.
R       100.1.1.0/24 [120/1] via 192.168.34.1, 00:27:01, FastEthernet 0/24
R       192.168.12.0/24 [120/2] via 192.168.34.1, 01:16:16, FastEthernet 0/24
R       192.168.23.0/24 [120/1] via 192.168.34.1, 01:16:16, FastEthernet 0/24
C       192.168.34.0/24 is directly connected, FastEthernet 0/24
C       192.168.34.2/32 is local host.
```

练习题

1. 选择题

（1）重发布到 RIP 域中路由的默认度量值是（　　）。

A．1　　B．10　　C．20　　D．30

（2）重发布到 OSPF 域中路由的默认度量值是（　　）。

A．1　　B．10　　C．20　　D．30

（3）OSPF 路由对 RIP 路由进行重发布，这些重发布路由的初始 RIP 度量值设置为 3，正确的配置命令是（　　）。

A．ruijie(config)#router rip
ruijie(config-router)#default-metric 3
ruijie(config-router)#redistribute ospf 100

B．ruijie(config)#router rip
ruijie(config)#default-metric 3
ruijie(config)#redistribute ospf 100

C．ruijie(config)#router ospf 1
ruijie(config-router)#default-metric 3
ruijie(config-router)#redistribute RIP

D．ruijie(config)#router rip
ruijie(config-router)#redistribute ospf 100

（4）在重发布到 OSPF 路由时，“redistribute static metric 30 subnets”命令表示（　　）。

A．禁止标准静态路由重发布到 RIP
B．禁止所有路由重发布 RIP
C．允许将所有路由重发布到 RIP
D．允许将非标准静态路由重发布到 RIP

（5）在重发布默认路由时，“default-information originate always”命令表示（　　）。

A．禁止重发布默认路由
B．禁止发布所有默认路由
C．无条件重发布默认路由
D．允许发布默认路由，且默认度量值为 20

2. 简答题

（1）描述路由重发布的工作原理。

（2）重发布到 RIP 中时，静态路由和直连路由的默认度量值是多少？其他重发布路由的默认度量值是多少？

（3）重发布到 OSPF 中时，静态路由和直连路由的默认度量值是多少？其他重发布路由的默认度量值是多少？

（4）说明重发布的配置命令及步骤。

项目 11　策略路由

在常规路由选择规则中，路由器取出报文中的目标 IP 地址，通过查找路由表来决定报文转发路径。有时，为了便于管理网络，需要根据管理者要求和策略来操纵报文的转发路径。使用 PBR（Policy-Based Routing，基于策略的路由）可以实现报文有条件地路由转发。

本章主要介绍路由器策略路由相关知识（包括定义及应用），通过项目实训掌握路由器策略路由的配置技能。

知识点、技能点

1. 了解路由器策略路由实现功能;
2. 掌握路由器策略路由配置方法。

11.1　问题提出

在网络应用中，有时需要根据数据包源 IP 地址或根据源地址、目标地址、接口号、协议及数据包长度控制数据包转发路径。如何满足这种需求呢？可以通过配置策略路由实现路由器有条件地转发数据包。

11.2　相关知识

1. 策略路由概述

（1）路由表产生。

路由器通过手工或路由协议获得路由表。手工设置路由是管理员依据网络结构确定路由走向手工设置路由，具体简单、成本低，适合小型网络；动态路由协议（如 RIP 协议或 OSPF 协议）通过交换路由信息、执行一定路由算法，选择最优路径产生路由表，能动态地实时反映网络拓扑变化，适合大中型网络。

如果一台路由器同时实现多个不同路由协议，不同路由协议获得达到同一网络的不同路径路由信息，路由器将比较不同路由协议的管理距离值，管理距离值越小，可信度越高，RIP 默认值为 120，OSPF 默认值为 110。路由器优先将管理距离值小的路由放入路由表。

在同一种路由协议中，如何评价不同路径的优劣呢？通过度量值来表示路径的好坏，如 RIP 以跳数来衡量；OSPF 以成本开销 Cost 来衡量。度量值越小，路径越优。如到达某网络的度量值一条为 10，另一条为 50，则优先选择度量值低的。

总之，路由表中的路由都是经过比较管理距离值、比较度量值后，选择的最优路由。

（2）传统路由转发。

当路由器接收到报文后，路由器会查看 IP 报头中的目标地址信息，并开始对路由表进行查找，以查找是否有到达目标网络的路由。如果本地路由表包含到达目标网络的路由条目，那么路由器就会根据路由转发路径进行报文转发。如果路由表中没有匹配的网络，则查看路由表是否设置了默认路由，

有默认路由则按默认路由转发，没有默认路由则直接丢弃该报文。

如果路由器接收到的报文的目的地是 172.16.2.26，但本地路由表中有两条到达该目的地的路由：172.16.0.0/16 和 172.16.2.0/24，但这个路由条目的掩码长度不一致。在这种情况下，路由器会根据最长掩码匹配的原则（即最精确匹配）选择掩码长度为 24 位的路由条目 172.16.2.0/24 来转发报文，如图 11.1 所示。

```
Router#show ip route

Codes: C - connecter, S - static, R - RIP B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default

Cateway of last resort is no set
R    172.16.0.0/16 [120/1] via 192.168.1.2, 00:23:31, FastEthernet 1/0
R    172.16.0.0/24 [120/1] via 192.168.3.2, 00:23:04, FastEthernet 1/1
C    192.168.1.0/24 is directly connected, FastEthernet 1/0
C    192.168.1.1/32 is local host.
C    192.168.3.0/24 is directly connected, FastEthernet 1/1
C    192.168.3.1/32 is local host.
```

图 11.1　最长掩码匹配

（3）策略路由原理。

策略路由 PBR 为网络管理者提供了比传统的路由选择对报文更强的控制能力。在当今的互连网络中，常常需要按照自己定义的策略，以超越传统路由选择的方式自由地实现报文转发和路径选择。基于策略的路由比传统的路由方式提供更强的操控性和灵活性，通过使用 PBR，管理员能够根据报文的源地址、目的地址、源接口、目的接口、协议类型、报文大小等为报文选择不同的转发路径。

当使用 PBR 时，路由器接收到报文后，首先根据预先制定的策略对报文进行匹配，如果匹配到一条策略，就根据该策略指定的方式对报文进行转发；如果没有匹配到任何策略，路由器将按照传统的路由选择方式，依据路由表中的条目对报文进行路由。

如图 11.2 所示，边界路由器作为企业网络的出口，通过两条链路连接到两个 ISP。图中也显示了路由器的部分路由表信息，可以看出路由器拥有两条静态默认路由，下一跳分别为 ISP1 和 ISP2。正常情况下，路由器将基于目的地址在这两条链路上对数据进行负载分担，并且我们无法控制报文的传输路径，这都是由路由器动态决定的。

```
S*   0.0.0.0/0 [1/0] via 192.0.1.2
S*             [1/0] via 192.0.2.2
C    192.0.1.0/30 is directly connected, Serial 0/0
C    192.0.1.0/32 is local host.
C    192.0.2.0/30 is directly connected, Serial 0/1
C    192.0.2.1/32 is local host.
```

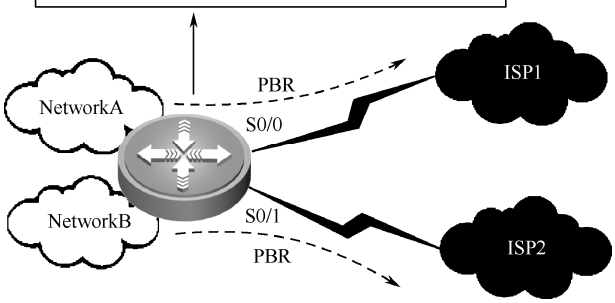


图 11.2　策略路由原理

如果我们希望企业网络中来自 NetworkA 的数据流被发送到 ISP1，而来自 NetworkB 的数据流被发送到 ISP2，则可以使用 PBR 来实现。在 PBR 中，我们可以为路由器设置策略，使其将收到的源地址为 NetworkA 的报文发送到 S0/0 链路上，同样将收到的源地址为 NetworkB 的报文发送到 S0/1 链路上，这样就达到了基于策略进行路由选择的目的。

在网络中使用 PBR 可以提供如下优点。

（1）灵活地选择路径：可以使用 PBR 根据不同的需求将不同用户或网络产生的数据流通过不同的路径传输。

（2）提高服务质量：可以使用 PBR 在网络边缘为报文设置不同的优先级，然后在网络的核心中利用各种 QoS（Quality of Service）机制对不同优先级的数据提供不同的服务等级，这样可以改善网络的性能。

（3）节省费用：可以让与特定业务相关的较大流量使用一条高带宽、高费用的链路，同时在一条低带宽、低费用的链路上为交互式数据流提供服务，从而合理地利用链路带宽。

（4）负载均衡：除了基于目标地址路由选择所提供的动态负载均衡功能外，通过 PBR 还可以实现多条路径之间的流量负载均衡。

PBR 是一种入站机制，用于在路由器接入接口对收到的报文进行处理。PBR 根据预先制定的策略将报文转发到下一跳地址或本地出接口。PBR 的操作优先于路由器查找路由表操作。网络管理员能够根据下列准则制定路由选择策略：源系统的身份；使用的协议；报文的长度。

2. 配置命令

（1）创建 route-map 命令。

在路由器上实现基于策略路由 PBR 需要使用一个非常重要的工具，它就是路由映射表（route-map）。route-map 用来实现报文的条件匹配和报文路由规则的设置。只有当报文匹配 route-map 中定义的特征或条件时，才会对报文进行基于策略的路由选择。

配置策略路由的第一步就是创建 route-map。一个 route-map 由多条策略语句组成，每条策略语句都有不同的序号，在操作的时候按照序号从小到大顺序地检查和操作。每个策略语句都定义了一个或多个的匹配条件和路由操作行为。

创建 route-map 的命令格式如下：

定义的路由图名称　　*路由图策略语句序号*

ruijie(config)#route-map *route-map-name* **[permit | deny]** *[sequence-number]*

设备名称为ruijie 处于全局配置模式　　*指定符合match条件，是否执行set命令*

其中，*route-map-name* 为定义的路由图名称。

permit 表示若符合 match 定义的匹配条件时，按照 set 命令指定的操作将对数据包进行转发控制。若不符合 match 定义的匹配条件，则进入第二个路由图策略语句，直到最终执行了 set 命令。

deny 表示若符合 match 定义的匹配条件时，则不会执行任何 set 命令，直接查询路由表，按路由表转发。若不符合 match 定义的匹配条件，则将进入下一个路由图策略语句，直到最终执行了 set 命令。

sequence-number 为路由图策略对应的序号。低序号的策略优先得到使用，因此需要注意序号的设置。

默认情况下，没有配置路由图。使用该命令的 no 形式删除指定路由图的定义，如 ruijie(config)# no route-map redrip permit 10。

例如：创建路由图 set-pref，对于从 EBGP 接收的 2.0.0.0 网络的路由设置本地优先级为 200，其他

路由设置为默认值。

```
ruijie(config)#route-map set_pref permit 10
ruijie(config-route-map)#match ip addres 1
ruijie(config-route-map)#set local-preference 200
ruijie(config-route-map)#exit
ruijie(config)#route-map set_pref permit 20
ruijie(config)#access-list 1 permit 2.0.0.0 0.255.255.255
```

例如：将 OSPF 路由协议重发布 RIP 路由，要求只重发布跳数为 4 的 RIP 路由，在 OSPF 路由域中，该路由的类型为外部路由 type-1，初始度量值为 40，路由标记值设为 40。

```
ruijie(config)#router ospf
ruijie(config-router)#redistribute rip subnets route-map redrip
ruijie(config-router)#network 192.168.12.0 0.0.0.255 area 0
rujijie(config-router)#exit
ruijie(config)#route-map redrip permit 10
ruijie(config-route-map)#match metric 4
ruijie(config-route-map)#set metric 40
ruijie(config-route-map)#set metric-type type-1
ruijie(config-route-map)#set tag 40
```

（2）match 命令。

route-map 中的 match 命令用来设置报文进行策略路由的匹配条件。一个 route-map 子句中可以配置多条 match 命令，即设置多个匹配条件。

定义报文 ACL 过滤条件命令为 match ip address。match ip address 命令通过标准访问列表或扩展访问列表定义策略路由条件，符合定义的条件，如果设置了 permit，则执行后面的 set 命令；如果定义 deny，则不执行 set 命令，直接执行传统路由。不符合定义的条件，继续查询下一个 match 语句。

match ip address 命令格式如下：

```
                                  定义的ACL过滤条件                                  多个ACL名称
ruijie(config-route-map)# match ip address { access-list-number | name } [ ...access-list-number | ...name ]
设备名称为ruijie                                  ACL名称
处于路由图配置模式
```

其中，*access-list-number* 为访问列表编号，标准访问列表编号范围为 1～99，1300～1999；扩展访问列表编号范围为 100～199，2000～2699。*name* 为访问列表名称。

例如：将 OSPF 路由协议重发布 RIP 路由，要求只重发布符合访问列表 10 的 RIP 路由，在 OSPF 路由域中，该路由的类型为外部路由 type-1，初始度量值为 40。

```
ruijie(config)#router ospf
ruijie(config-router)#redistribute rip subnets route-map redrip
ruijie(config-router)#network 192.168.12.0 0.0.0.255 area 0
ruijie(config-router)#exit
ruijie(config)#access-list 10 permit 200.168.23.0 0.0.0.255
ruijie(config)#route-map redrip permit 10
ruijie(config-route-map)#match ip address 10
ruijie(config-route-map)#set metric 40
ruijie(config-route-map)#set metric-type type-1
```

（3）set 命令。

如果满足匹配语句（match 命令）的条件，则可以使用一个或多个 set 命令来指定对报文的操作，即如何对报文执行基于策略的路由。通常我们都是为满足匹配条件的报文指定下一跳地址或出接口。

当使用 set 命令设置路由策略时，有些命令只影响那些在路由表中存在相应的显式路由的报文，而一些命令只影响在路由表中不存在的显式路由的报文。

需要注意的是，对于将操作设置为 deny 的 route-map，如果报文匹配 match 命令设置的条件，报文将不被执行策略路由，而是按照传统的方式进行路由，即使配置了 set 命令，它也不会被执行。

① 配置下一跳地址的命令。为了对于满足 match 条件的报文按照指定的下一跳地址进行转发，使用如下命令设置下一跳地址：

指定报文转发的下一跳地址

ruijie(config-route-map)# set ip next-hop *ip-address [...ip-address]*

设备名称为ruijie
处于路由图配置模式　　*下一跳地址*

其中，*ip-address* 为下一跳 IP 地址。如果指定了多个下一跳地址，当与第一个下一跳地址相关联的本地接口的状态为 Down 时，路由器会按照顺序轮流地尝试后续的下一跳地址。

使用 set ip next-hop 命令时，路由器将检查路由表以确定是否可以到达下一跳地址，而不会检查是否存在前往报文目标地址的显式路由。

例如：在 Serial 1/0 接口上启用策略路由。当该接口接收到源地址在 10.0.0.0/8 范围内的数据包的流量，将发送到 192.168.100.1；源网络为 172.16.0.0/16 范围内的数据包的流量，将发送到 172.16.100.1；其余的数据流量将全部丢弃。

```
ruijie(config)#interface serial 1/0
ruijie(config-if)#ip policy route-map load-balance
ruijie(config)#access-list 10 permit 10.0.0.0    0.255.255.255
ruijie(config)#access-list 20 permit 172.16.0.0    0.0.255.255
ruijie(config)#route-map load-balance permit 10
ruijie(config-route-map)#match ip address 10
ruijie(config-route-map)#set ip next-hop 192.168.100.1
ruijie(config-route-map)#exit
ruijie(config)#route-map load-balance permit 20
ruijie(config-route-map)#match ip address 20
ruijie(config-route-map)#set ip next-hop 172.16.100.1
ruijie(config-route-map)#exit
ruijie(config)#route-map load-balance permit 30
ruijie(config-route-map)#set interface Null 0
```

② 配置转出接口的命令。为了对于满足 match 条件的报文按照指定的接口转发，使用如下命令设置转发接口：

指定报文转发接口　　*接口编号*

ruijie(config-route-map)# set interface *interface-type interface-number [...interface-type interface-number]*

设备名称为ruijie
处于路由图配置模式　　*接口类型*

其中，*interface-type* 为接口类型，如 Fastethernet、Serial 等；*interface-number* 为接口编号，如 0/0、2/0 等。

route-map 中的配置命令 set interface interface […interface] 提供了一个接口列表用于指定符合匹配条件的报文被转发的本地转出接口。如果指定了多个转出接口，第一个状态为 Up 的接口将用于转发报文，如果接口状态为 Down，路由器将尝试后续的转出接口。

如果要将符合匹配条件的报文丢弃，而不是进行传统的路由选择，可以使用 set interface null 0 命令将报文发送到 null 0 接口。null 0 接口是路由器中的一个虚拟的空接口，发送到该接口的报文将被丢弃。

例如：在 Fastethernet 0/0 接口启用策略路由，当该接口接收到数据包小于 500 字节的流量，将从 Serial 2/0 接口发送数据包。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip policy route-map smallpak
ruijie(config)#route-map smallpak permit 10
ruijie(config-route-map)#match length 0 500
ruijie(config-route-map)#set interface serial 2/0
ruijie(config-route-map)#exit
```

（4）route-map 操作原则。

在使用 route-map name [permit | deny] [sequence-number] 命令创建 route-map 时，名称（name）相同而序号（sequence-number）不同的多个 route-map 子句将组成一个 route-map，在执行 route-map 时将按照序号从小到达的顺序执行。当执行到某条子句并发生匹配时（满足 match 命令的匹配条件），如果 route-map 的操作为 permit，则执行 set 命令，并且不执行后续的 route-map 子句；如果 route-map 的操作为 deny，则不执行 set 命令和后续的 route-map 子句。

route-map 中的 permit 和 deny 并不表示允许和拒绝报文通过，这与访问控制列表不同。route-map 中的允许和拒绝表示是否对符合匹配条件的报文应用策略。

route-map 与访问控制列表一样，默认的操作都为 deny，即在每个 route-map 的最后都会隐含着一个 deny any 的子句。如果在所有配置的 route-map 子句中都没有发生匹配，路由器将执行最后的 deny 子句，对于 PBR 来说，就是不对报文进行策略路由，而是按照传统的方式路由，如图 11.3 所示。

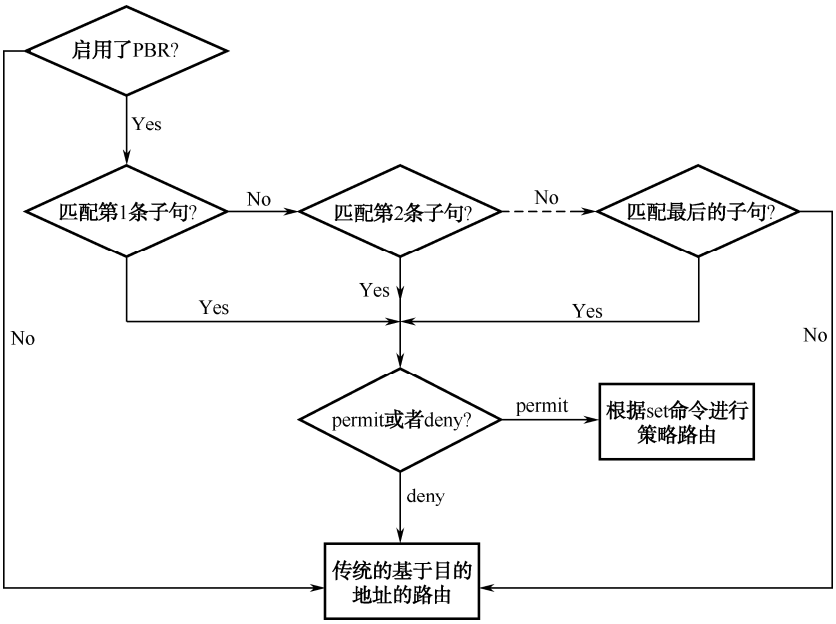


图 11.3　route-map 操作流程

在一个 route-map 子句中，可以配置多个 match 命令和 set 命令。一个 route-map 子句中的多个 match 命令表示的关系是“and”，一个 match 命令中的多个条件表示的关系是“or”，set 命令也是如此。但如果同时设置了下一跳和转出接口，路由器将只会执行第一个 set 命令。如果同时设置了下一跳和 IP 优先级，那么两个 set 命令都会被执行。

例如：路由器策略路由配置。

```
ruijie(config)#route-map ruijie-pbr permit 10
ruijie(config-route-map)#match ip address 110 111 112
ruijie(config-route-map)#set ip next-hop 192.168.2.2
ruijie(config-route-map)#exit
ruijie(config)#route-map ruijie-pbr permit 20
ruijie(config-route-map)#match ip address 120
ruijie(config-route-map)#match ip address 121
ruijie(config-route-map)#set ip next-hop 192.168.2.6
ruijie(config-route-map)#exit
ruijie(config)#route-map ruijie-pbr permit 30
ruijie(config-route-map)#set ip next-hop 192.168.2.13
ruijie(config-route-map)#exit
```

在上述配置代码中，名称为 ruijie-pbr 的 route-map 包括三条子句，序号分别为 10、20、30，路由器将按照序号从小到大的顺序进行匹配。

序号为 10 的子句表示，如果报文符合访问控制列表 110、111 和 112 三个中的任何一个，则匹配成功，并执行 set 命令，将报文发送到下一跳地址 192.168.2.2。如果报文不匹配其中任何一个访问控制列表，路由器将执行序号为 20 的子句。在该子句中，如果报文同时匹配访问控制列表 120 和 130，则匹配成功，并执行 set 命令，将报文发送到下一跳地址 192.168.2.6。如果报文不是同时匹配这两个访问控制列表，则执行序号为 30 的子句。在序号为 30 的子句中，没有配置 match 命令，这表示默认匹配所有，并将不匹配 10 和 20 子句的所有报文发送到下一跳地址 192.168.2.13。

如果上述配置代码中没有序号为 30 的子句，则执行隐含的 deny any，将不匹配 10 和 20 子句的所有报文进行传统的路由选择。

最后，当我们要删除 route-map 时，如果不指定具体的序号，则将会删除整个 route-map。

（5）应用路由策略。

完成 route-map 的配置后，要使路由器能够对接收的报文执行基于策略的路由选择，需要在报文的入接口上应用 route-map。

在接口下使用命令 ip policy route-map name 将 route-map 应用到接口，此处的 name 要与之前配置的 route-map 的名称一致。PBR 是一个入站特性，所以需要在接收报文（而不是发送报文）的接口上配置。

将设置的路由策略应用到接口上的命令格式如下：

其中，*route-map* 为路由图名称。

例如：当快速以太网接口 Fa0/0 接收到数据报，如果数据报源地址为 10.0.0.1，则设置下一跳为 196.168.4.6，如果源地址为 20.0.0.1，则设置下一跳为 196.168.5.6，否则进行普通转发。

```
ruijie(config)#access-list 1 permit 10.0.0.1
ruijie(config)#access-list 2 permit 20.0.0.1
ruijie(config)#route-map lab1 permit 10
ruijie(config-route-map)#match ip address 1
ruijie(config-route-map)#set ip next-hop 196.168.4.6
ruijie(config-route-map)#exit
ruijie(config)#route-map lab1 permit 20
ruijie(config-route-map)#match ip address 2
ruijie(config-route-map)#set ip next-hop 196.168.5.6
ruijie(config-route-map)#exit
ruijie(config)#interface FastEthernet 0/0
ruijie(config-if)#ip policy route-map lab1
ruijie(config-if)#exit
```

11.3 扩展知识

1. Match 语句中定义报文长度过滤条件命令

不仅可以通过引用访问控制列表根据地址、接口等信息来匹配报文，还可以根据报文长度进行策略路由。要配置基于报文长度的匹配条件时可以使用如下命令：

定义的报文长度过滤条件　　*报文最大长度*

ruijie(config-route-map)#match length *min-length max-length*

设备名称为ruijie
处于路由图配置模式　　*报文最小长度*

其中，*min-lengh* 为报文最小长度；*max-length* 为报文最大长度。

可以将匹配报文长度作为区分交互数据流和文件传送数据流的条件，因为文件传送数据报文通常较大，交互数据流的报文通常较小。

例如：在 Fastethernet 0/0 接口启用策略路由，将数据包小于 500 字节的流量，从 Serial 2/0 接口发送出去。

```
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip policy route-map smallpak
ruijie(config-if)#exit
ruijie(config)#route-map smallpak permit 10
ruijie(config-route-map)#match length 0 500
ruijie(config-route-map)#set interface serial 2/0
ruijie(config-route-map)#exit
```

2. Set 语句中配置默认下一跳地址命令

为了对报文提供一个默认下一跳地址，使用如下命令设置默认下一跳地址：

指定报文默认下一跳地址

ruijie(config-route-map)# set ip default next-hop *ip-address […ip-address]*

设备名称为ruijie
处于路由图配置模式

默认下一跳地址

其中，*ip-address* 为下一跳 IP 地址。route-map 中的配置命令 set ip default next-hop ip-address […ip-address] 提供了一个默认下一跳地址列表。所谓默认下一跳是指当路由表中没有到达报文目的地址的显式路由时，才将报文发送到指定的默认下一跳地址。如果指定了多个默认下一跳地址，当与第一个默认下一跳地址相关联的本地接口的状态为 Down 时，路由器会按照顺序轮流地尝试后续的默认下一跳地址。

set ip next-hop 命令和 set ip default next-hop 命令都是为报文指定传输的下一跳路径，但是它们的不同点是：使用 set ip next-hop 命令将使路由器首先进行基于策略的路由选择，然后再进行查找路由表的传统的路由选择；相反，使用 set ip default next-hop 命令时，路由器将先使用传统的路由选择方式，如果无法找到匹配的路由条目，才会执行基于策略的路由选择。

例如：对两个不同的节点发出的报文分别转发到不同的路径上去。在 Fastethernet 0/0 接口上接收的从 1.1.1.1 发出的报文，如果路由器找不到转发路由时，就被转发到设备 6.6.6.6 去；而接收到的从 2.2.2.2 发出的报文，如果路由器找不到转发路由时，就被转发到设备 7.7.7.7 去。其他的报文如果路由器找不到转发路由时就丢弃处理。

```
ruijie(config)#access-list 1 permit ip 1.1.1.1 0.0.0.0
ruijie(config)#access-list 2 permit ip 2.2.2.2 0.0.0.0
ruijie(config)#interface fastethernet 0/0
ruijie(config-if)#ip policy route-map equal-access
ruijie(config-if)#exit
ruijie(config)#route-map equal-access permit 10
ruijie(config-route-map)#match ip address 1
ruijie(config-route-map)#set ip default next-hop 6.6.6.6
ruijie(config-route-map)#exit
ruijie(config)#route-map equal-access permit 20
ruijie(config-route-map)#match ip address 2
ruijie(config-route-map)#set ip default next-hop 7.7.7.7
ruijie(config-route-map)#exit
ruijie(config)#route-map equal-access permit 30
ruijie(config-route-map)#set default interface null 0
ruijie(config-route-map)#exit
```

3. Set 语句中配置默认转出接口命令

为了对报文提供一个默认转出接口，可以使用如下命令设置默认转出接口：

指定报文默认转出接口　　转出接口编号

ruijie(config-route-map)#set default interface *interface-type interface-number […interface-type interface-number]*

设备名称为ruijie
处于路由图配置模式　　接口类型

其中，*interface-type* 为接口类型；*interface-number* 为接口编号。

与设置默认下一跳一样，我们使用 route-map 中的配置命令 set default interface interface […interface]

设置一个默认的转出接口列表。当路由表中没有前往其目的地址的显式路由时，报文将被路由到指定的默认转出接口。如果指定了多个默认转出接口，第一个状态为 Up 的接口将用于转发报文，如果接口状态为 Down，路由器将尝试后续的默认转出接口。

与设置下一跳和默认下一跳一样，使用 set interface 命令将使路由器首先进行基于策略的路由选择，然后再进行查找路由表的传统的路由选择；相反，使用 set default interface 命令将使路由器先执行传统的路由选择，如果无法找到匹配的路由条目，才会执行基于策略的路由选择。

例如：在 Fastethernet 0/0 接口启用策略路由，当该接口接收到数据包小于 500 字节的流量，而且路由表中没有明确的路由时，将从 Serial 2/0 接口发送数据包。

```
ruijie(config)#interface    fastethernet 0/0
ruijie(config-if)#ip policy route-map smallpak
ruijie(config)#route-map smallpak permit 10
ruijie(config-route-map)#match length 0 500
ruijie(config-route-map)#set default interface serial 2/0
ruijie(config-route-map)#exit
```

11.4 演示实例

1. 背景描述

某单位计算机网络出口采用双出口接入外网，其中，一个出口接 Cernet 中国教育和科研网，另一个出口接 Chinanet。现需要对边界路由器进行配置，实现访问 Cernet 的报文从 S2/0 接口发送，其他流量从 S3/0 接口发送，网络拓扑结构如图 11.4 所示。

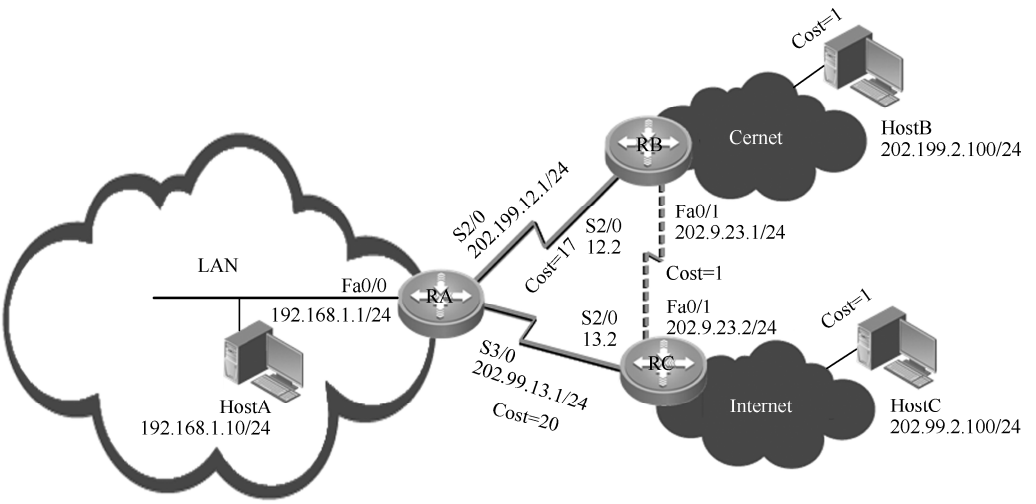


图 11.4　策略路由应用

2. 操作步骤

若要对访问 Cernet 网络的流量进行过滤，就必须知道 Cernet 的地址范围，然后通过设置 ACL，实现对相关流量进行过滤的功能。Cernet 网络地址为 202.199.0.0 /24。

设备参数配置如下。

第 1 步：配置路由器 IP 地址。

```
ruijie(config)#hostname RA
RA(config)#interface Fastethernet 0/0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface serial 2/0
RA(config-if)#ip address 202.199.12.1    255.255.255.0
RA(config-if)#clock rate 64000
RA(config-if)#bandwidth 6000
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#interface serial 3/0
RA(config-if)#ip address 202.99.13.1 255.255.255.0
RA(config-if)#clock rate 64000
RA(config-if)#no shutdown
RA(config-if)#exit
RA(config)#
```

```
ruijie(config)#hostname RB
RB(config)#interface Fastethernet 0/0
RB(config-if)#ip address 202.199.2.1 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
RB(config)#interface Fastethernet 0/1
RB(config-if)#ip address 202.9.23.1 255.255.255.0
RB(config-if)#no shutdown
RB(config-if)#exit
RB(config)#interface serial 2/0
RB(config-if)#ip address 202.199.12.1    255.255.255.0
RB(config-if)#bandwidth 6000
RB(config-if)#no shutdown
RB(config-if)#exit
```

```
ruijie(config)#hostname RC
RC(config)#interface Fastethernet 0/0
RC(config-if)#ip address 202.99.2.1 255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
RC(config)#interface Fastethernet 0/1
RC(config-if)#ip address 202.9.23.2 255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
RC(config)#interface serial 2/0
```

```
RC(config-if)#ip address 202.99.13.1    255.255.255.0
RC(config-if)#no shutdown
RC(config-if)#exit
```

第 2 步：配置 OSPF 路由协议。

```
RA(config)#router ospf 1
RA(config-router)#network 192.168.1.0 0.0.0.255 area 10
RA(config-router)#network 202.199.12.0 0.0.0.255 area 0
RA(config-router)#network 202.99.13.0 0.0.0.255 area 0
RA(config-router)#exit
RA(config)#
```

```
RB(config)#router ospf 1
RB(config-router)#network 202.199.2.0 0.0.0.255 area 20
RB(config-router)#network 202.199.12.0 0.0.0.255 area 0
RB(config-router)#network 202.9.23.0 0.0.0.255 area 0
RB(config-router)#exit
RB(config)#
```

```
RC(config)#router ospf 1
RC(config-router)#network 202.99.2.0 0.0.0.255 area 30
RC(config-router)#network 202.99.13.0 0.0.0.255 area 0
RC(config-router)#network 202.9.23.0 0.0.0.255 area 0
RC(config-router)#exit
RC(config)#
```

第 3 步：查看路由表。

```
RA#show ip route
Codes:      C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       192.168.1.0/24 is directly connected, FastEthernet 0/0
C       192.168.1.1/32 is local host.
O       202.9.23.0/24 [110/18] via 202.199.12.2, 00:32:14, Serial 2/0
O IA 202.99.2.0/24 [110/19] via 202.199.12.2, 00:32:14, Serial 2/0
C       202.99.13.0/24 is directly connected, Serial 3/0
C       202.99.13.1/32 is local host.
O IA 202.199.2.0/24 [110/18] via 202.199.12.2, 00:25:18, Serial 2/0
C       202.199.12.0/24 is directly connected, Serial 2/0
C       202.199.12.1/32 is local host.
```

```
RB#show ip route
Codes:    C - connected, S - static, R - RIP, B - BGP
      O - OSPF, IA - OSPF inter area
      N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
      E1 - OSPF external type 1, E2 - OSPF external type 2
      i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
      ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
O IA 192.168.1.0/24 [110/18] via 202.199.12.1, 00:31:42, Serial 2/0
C       202.9.23.0/24 is directly connected, FastEthernet 0/1
C       202.9.23.1/32 is local host.
O IA 202.99.2.0/24 [110/2] via 202.9.23.2, 00:44:35, FastEthernet 0/1
O       202.99.13.0/24 [110/49] via 202.9.23.2, 00:44:35, FastEthernet 0/1
C       202.199.2.0/24 is directly connected, FastEthernet 0/0
C       202.199.2.1/32 is local host.
C       202.199.12.0/24 is directly connected, Serial 2/0
```

```
RC#show ip route
Codes: C - connected, S - static,    R - RIP
          O - OSPF, IA - OSPF inter area
          E1 - OSPF external type 1, E2 - OSPF external type 2
Gateway of last resort is not set
O       202.199.12.0/24 [110/18] via 202.9.23.1, 00:26:34, FastEthernet1
O IA 202.199.2.0/24 [110/2] via 202.9.23.1, 00:26:34, FastEthernet1
O IA 192.168.1.0/24 [110/19] via 202.9.23.1, 00:26:34, FastEthernet1
C       202.99.13.0/24 is directly connected, Serial2/0
C       202.99.2.0/24 is directly connected, FastEthernet0
C       202.9.23.0/24 is directly connected, FastEthernet1
```

从以上路由表可以看出，去往 Cernet 网络及去往 Internet 的所有数据流量都经过 RB 路由器。原因是 OSPF 路由协议的度量值 Cost 是根据链路带宽衡量优劣的。

默认情况下，OSPF 接口 Cost 值为 100Mbps/Bandwidth。路由器 Serial 接口默认 Bandwith 为 2000Kbit，Fastethernet 接口默认 Bandwith 为 100Mbit。在这个例子中我们将 Serial 2/0 的 Bandwith 设置为 6000Kbit，故 S2/0 的 Cost 为 17，S3/0 为 20，Fa0/0 为 1 等。

如图 11.5 所示网络，去往 202.199.2.0 网络的数据流量经过 RA→RB 路径，去往 202.99.2.0 网络的数据流量要经过 RA→RB→RC 路径。

我们通过对 RA 设置策略路由，改变去往 202.199.2.0 网络的数据流量行走路径，使其通过 RA→RC 直接到达 202.199.2.0 网络。具体参数配置如下。

第 4 步：配置策略路由。

```
RA(config)#access-list 100 permit ip any 202.199.0.0 0.0.255.255
RA(config)#route-map cernet permit 10
RA(config-route-map)#match ip address 100
RA(config-route-map)#set next-hop 202.199.12.2
RA(config-route-map)#exit
```

```
RA(config)#route-map cernet permit 20
RA(config-route-map)#set net-hop 202.99.13.2
RA(config-route-map)#exit
RA(config)#
```

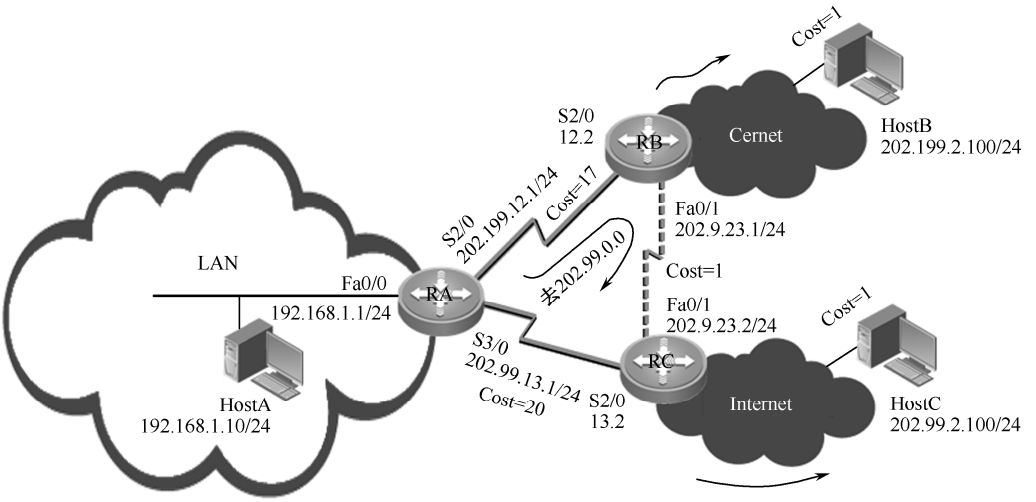


图 11.5　未应用策略路由情况

第 5 步：应用策略路由。

```
RA(config)#interface fastethernet 0/0
RA(config-if)#ip policy route-map cernet
RA(config-if)#
```

经过上述配置，使得去往 202.99.2.0 网络的数据流量通过路由器 RA→RC 到达，如图 11.6 所示。

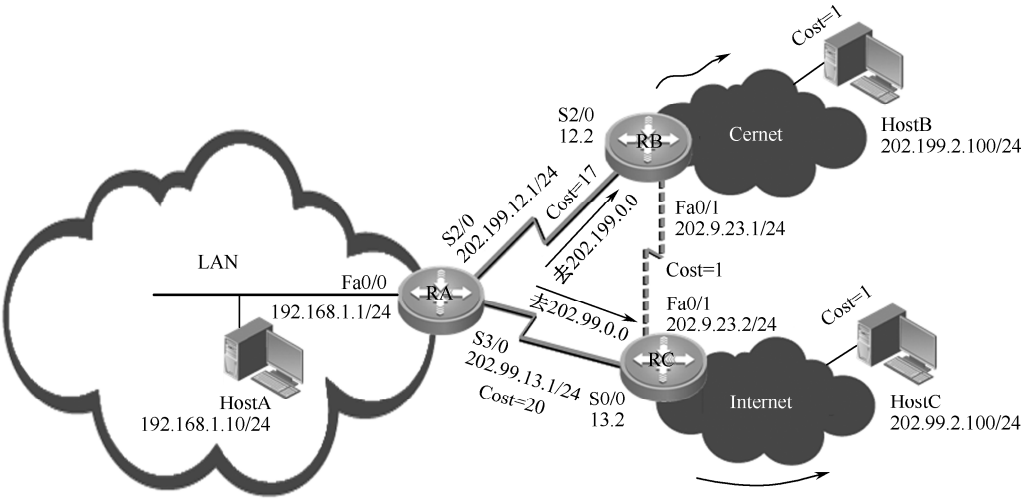


图 11.6　应用策略路由情况

第 6 步：测试验证。

在主机 A 的命令行下执行如下命令可以观察到数据包的行走路径。

```
C:\>tracert 202.199.2.100
```

```
Tracing route to 202.99.2.100 over a maximum of 30 hops:
1     <1ms     <1ms     <1ms     192.168.1.1
2     26ms     25ms     25ms     202.99.13.2
3     30ms     30ms     30ms     202.99.2.100
```

11.5　训练任务

1. 背景描述

在如图 11.7 所示网络中，所有网络都使用 192.168.0.0/24 设置 IP 地址，如 PC1 的 IP 地址为 192.168.1.10/24。在 SW1 三层交换机上配置策略路由，实现 PC1 访问 PC2 时经过 SW1→R2 路径；PC1 访问 PC3 时经过 SW1→R3 路径；其他流量均丢弃。SW1、R2 及 R3 都启用 OSPF 动态路由协议。

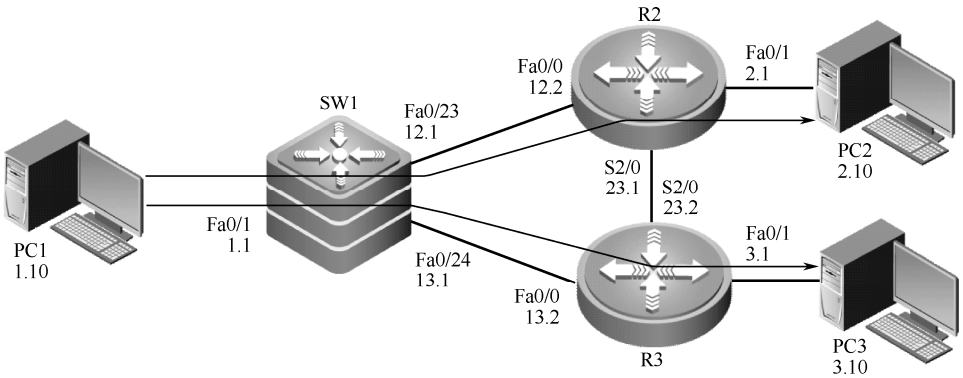


图 11.7　策略路由技能训练

2. 操作提示

（1）使用 no switchport 命令将三层交换机 Fa0/1、Fa0/23 及 Fa0/24 设置为路由接口，并相应设置 IP 地址。

```
SW1(config)#interface fastethernet 0/1
SW1(config-if)#no switchport
SW1(config-if)#ip address 192.168.1.1 255.255.255.0
SW1(config-if)#no shutdown
SW1(config-if)#exit
SW1(config)#interface fastethernet 0/23
SW1(config-if)#no switchport
SW1(config-if)#ip address 192.168.12.1 255.255.255.0
SW1(config-if)#no shutdown
SW1(config)#interface fastethernet 0/24
SW1(config-if)#no switchport
SW1(config-if)#ip address 192.168.13.1 255.255.255.0
SW1(config-if)#no shutdown
```

（2）在 SW1、R2 及 R3 上配置 OSPF 动态路由协议。

由于前面已经详细介绍过，在此略。

（3）在 SW1 上配置策略路由。

```
SW1(config)#access-list 101 permit ip any host192.168.2.10
SW1(config)#access-list 102 permit ip any host192.168.3.10
SW1(config)#route-map pc1topc2 permit 10
SW1(config-route-map)#match ip address 101
SW1(config-route-map)#set next-hop 192.168.12.2
SW1(config-route-map)#exit
SW1(config)#route-map pc1topc2 permit 20
SW1(config-route-map)#match ip address 102
SW1(config-route-map)#set net-hop 192.168.13.2
SW1(config-route-map)#exit
SW1(config)#route-map pc1topc2 permit 30
SW1(config-route-map)#set interface null 0
SW1(config-route-map)#exit
SW1(config)#
```

（4）在 PC1 上分别使用 C:\>tracert 192.168.2.10 及 C:\>tracert 192.168.3.10 命令进行验证测量。

练习题

1. 选择题

（1）下面对策略路由的描述正确的是（　　）。

A．策略路由是在报文离开接口时实施的路由策略

B．策略路由只能根据报文来源进行路由控制

C．策略路由优先于查询路由表对录音进行控制

D．策略路由不根据报文长度决定路由走向

（2）下面配置代码中，match ip address 10 语句中的“10”代表（　　）。

```
ruijie(config)#route-map load-balance permit 10
ruijie(config-route-map)#match ip address 10
ruijie(config-route-map)#set ip next-hop 192.168.100.1
ruijie(config-route-map)#exit
ruijie(config)#route-map load-balance permit 20
ruijie(config-route-map)#match ip address 20
ruijie(config-route-map)#set ip next-hop 172.16.100.1
ruijie(config-route-map)#exit
```

A．路由策略语句序号

B．允许发生报文的大小

C．允许发生报文的数量

D．ACL 编号

（3）在下面配置代码中，set interface null 0 语句的作用是（　　）。

```
ruijie(config)#route-map equal-access permit 10
ruijie(config-route-map)#match ip address 1
ruijie(config-route-map)#set ip default next-hop 6.6.6.6
ruijie(config-route-map)#exit
```

```
ruijie(config)#route-map equal-access permit 20
ruijie(config-route-map)#match ip address 2
ruijie(config-route-map)#set ip default next-hop 7.7.7.7
ruijie(config-route-map)#exit
ruijie(config)#route-map equal-access permit 30
ruijie(config-route-map)#set default interface null 0
ruijie(config-route-map)#exit
```

A．将路由器接收到的所有报文都发送至 null 0

B．将不符合条件语句 match ip address 1 和 match ip address 2 的所有报文都发送至 null 0

C．将不符合条件语句 match ip address 1 和 match ip address 2，且在路由表中查询不到的所有报文都发送至 null 0

D．对所有报文都不做任何处理

（4）在下面配置代码中，110、111、112 之间的逻辑关系是（　　）。

```
ruijie(config)#route-map ruijie-pbr permit 10
ruijie(config-route-map)#match ip address 110 111 112
ruijie(config-route-map)#set ip next-hop 192.168.2.2
ruijie(config-route-map)#exit
```

A．与　　B．或　　C．异或　　D．取反

（5）在下面配置代码中，120、121 之间的逻辑关系是（　　）。

```
ruijie(config)#route-map ruijie-pbr permit 20
ruijie(config-route-map)#match ip address 120
ruijie(config-route-map)#match ip address 121
ruijie(config-route-map)#set ip next-hop 192.168.2.6
ruijie(config-route-map)#exit
```

A．与　　B．或　　C．异或　　D．取反

2. 简答题

（1）描述策略路由的工作过程。

（2）比较策略路由与常规路由，说明它们的异同点。

（3）说明策略路由的配置命令及步骤。

项目 12　BGP 路由协议与基本配置

在前面几个章节中学习的路由协议（RIP、OSPF 等）都是 IGP，它们都是在一个 AS 内部提供路由路径选择的，其中 RIP 是距离向量路由协议，OSPF 是链路状态路由协议。Internet 由多个 AS 集合，AS 内部由 IGP 协议提供路由，那么 AS 之间的路由是如何选择的呢？本章将介绍用于在 AS 之间提供路由选择的 BGP 路由协议。

知识点、技能点

1. 掌握 BGP 路由协议的特性及应用环境;
2. 掌握 BGP 的配置技能。

12.1　问题提出

某 ISP 由于业务增长迅速，网络规模需要扩大，并需要规划与其他 ISP 路由选择路径。您作为公司网络规划师负责规划、设计公司网络规模、出口流量控制策略、路由交换方式等。

12.2　相关知识

1. BGP 路由协议概述

（1）自治系统及编号。

通过前面的学习我们已经了解到 RIP、OSPF 都是 IGP 路由协议，它们都是在一个自治系统 AS（Autonomous System）内部提供选择路径机制的。所谓的自治系统 AS 是指一组共享相同的路由选择策略和管理策略，并在一个单独的管理域中运行的路由器集合。每个 AS 都是由专门的技术团队来管理和维护的，通常情况下，一个 ISP 就是一个 AS。整个 Internet 网络就是由若干个 AS 或 ISP 组成的。每个 AS 都有一个唯一的自治系统编号，该编号由 IANA（Internet Assigned Numbers Authority）进行统一分配。自治系统编号也称为 AS 号，是一个 16 位的二进制数，范围为 1～655365，其中 64512～65535 为私有 AS 号，可以在企业内部使用，其他非私有 AS 号在 Internet 中使用，如图 12.1 所示。

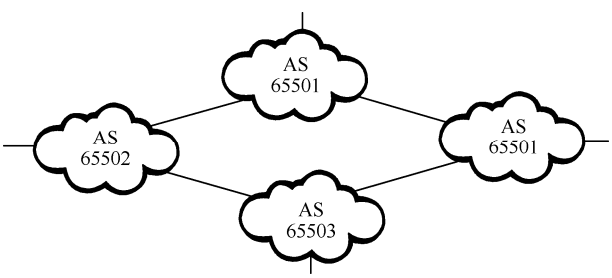


图 12.1　自治系统及编号

（2）IGP 与 EGP。

IGP 与 EGP 是路由协议的两大类型。在前面章节中学习的 RIP、OSPF 及 IS-IS 路由协议都是在一个 AS 内部运行的路由协议，为 AS 内部提供路由选择服务，我们将它们称为 IGP（Interior Gateway Protocol，内部网关协议）。与 IGP 相对应，有一种在不同自治系统之间提供路由选择机制的外部网关协议 EGP（Exterior Gateway Protocol，外部网关路由协议），EGP 运行在 AS 之间，如图 12.2 所示。

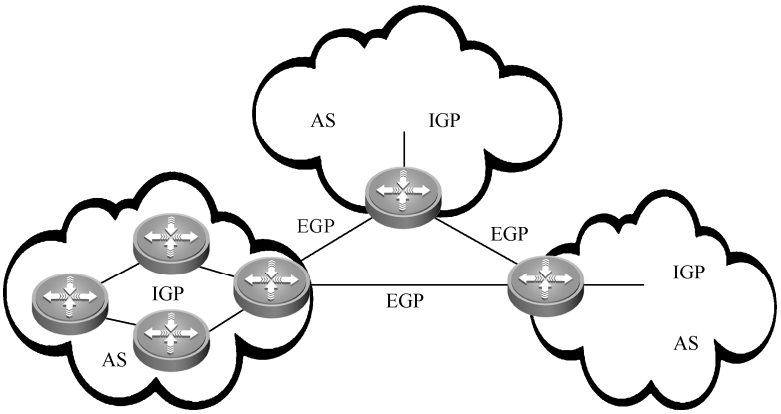


图 12.2　IGP 与 EGP 路由协议

最早的 EGP 协议是在 20 世纪 80 年代运行在 APPANET 中的，后来由于 EGP 不能检测环路，并且收敛速度慢、缺乏路由策略管理功能等，不能满足 Internet 高速发展需要而逐渐退出使用。

在 EGP 基础上开发的 BGP（Border Gateway Protocol）逐步承担了原来的 EGP 的使命。BGP 最初版本是在 RFC1105 中发布的，并且在 RFC1163 中得到了更新。随后 BGP 在 RFC1267 中被再一次更新。通常将这三个版本称为 BGP-1、BGP-2、BGP-3。目前在 Internet 上使用的是 BGP-4，出现在 RFC1771 中，它是一个无类路由选择协议。

每台运行 BGP 协议的路由器都会建立一个 BGP 数据库，存在本地 AS 和从其他 AS 接收的 BGP 路由信息，主要包括目的地址、下一跳、开销值、AS 路由列表等 BGP 路径属性。如果该路由器同时也运行了 IGP 协议（RIP、OSPF 等）时，那么该路由器还将存在 IP 路由表。当然 IP 路由表也可以包含 BGP 类型的路由信息。还可以通过路由重发布技术将 BGP 数据库的路由信息通告到 IGP 内，如图 12.3 所示。

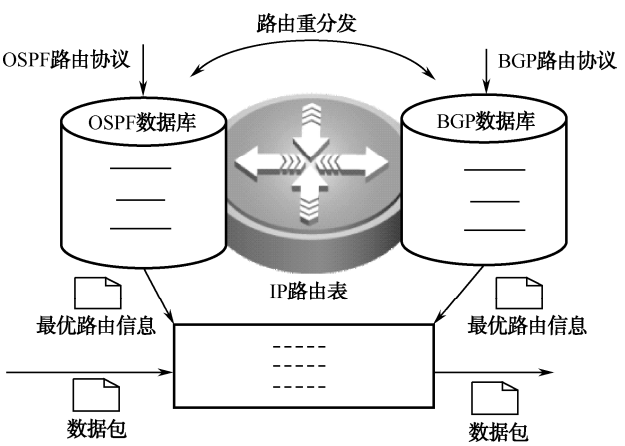


图 12.3　IP 路由表与 BGP 数据库

（3）BGP 主要特性。

BGP 是运行在 AS 之间的外部网关路由协议，以 AS 路径属性方式提供从一个 AS 到另一个 AS 的路由可达信息，有时也将 BGP 称为路径矢量路由协议，如图 12.4 所示。

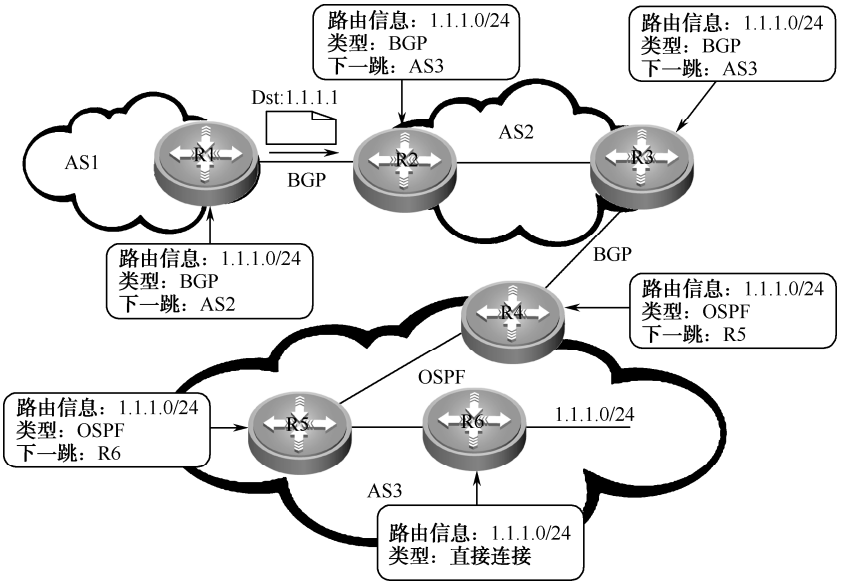


图 12.4　BGP 特性

在图 12.4 中，自治系统 AS1 中的路由器 R1 收到目的地址为 1.1.1.1 的数据包，它查看路由表中的 BGP 路由获知应将此数据包发往 AS2 中；AS2 中的路由器 R2 收到数据包后也通过查看路由表中的 BGP 路由获知应将此数据包发往 AS3 中，本 AS2 中的路由器 R3 与 AS3 相连接，所以应将此数据包发送至路由器 R3；路由器 R3 收到此数据包后通过查询路由表中的 BGP 路由获知应将此数据包发往 AS3 中的路由器 R4；路由器 R4 收到数据包后查询路由表，发现去往 1.1.1.0/24 的是一条类型为 OSPF 的 IGP 路由，并将此数据包发往路由器 R5；R5 收到此数据包后通过查询路由表后发往路由器 R6；最后由 R6 直接发送到目的地 1.1.1.1。

从以上数据包传递过程可以看出，BGP 是提供 AS 之间的路径可达信息，以 AS 作为“下一跳”，而不像 RIP 以路由器作为“下一跳”。BGP 衡量路径优劣不只是 AS 路径属性，还有其他很多参数，如优先级属性、起源属性、下一跳属性等，下面分别介绍。

第一，AS 路径列表。AS 路径列表是描述数据包为了到达目的地应走过的 AS 编号列表，也就是数据包在去往目的地的途中所走过的“路程”。BGP 路由器通告的路由更新报文中每一条路由都包含一个 AS 路径列表，每经过一个 AS 就将该路由器所在的 AS 号加入到这个列表中。在同一个 AS 内部传递路由更新报文时，AS 路径列表不变。

在如图 12.5 所示的网络中，AS1 中的 BGP 路由器将网络 192.168.1.0/24 路由更新报文通告给 AS2，这条 BGP 路由在离开 AS1 时附加了一个 AS 路径列表，由于网络 192.168.1.0/24 起源于 AS1，所以 AS 路径列表中只添加了 AS 编号，AS 路径列表为“1”。当 AS2 中的路由器再将此路由通告给 AS3 时，AS2 中的路由器将会把本地 AS 号添加到现有 AS 列表的起始位置，此时的 AS 列表为“2,1”。依次类推，当 AS5 从 AS4 及 AS6 收到这条 BGP 路由后，它看到的 AS 列表将是“4,3,2,1”和“6,7,3,2,1”。可以看到，AS 列表描述了这条路由所走过的“路程”，这也就是 BGP 为什么是一种“路径”矢量协议的原因。

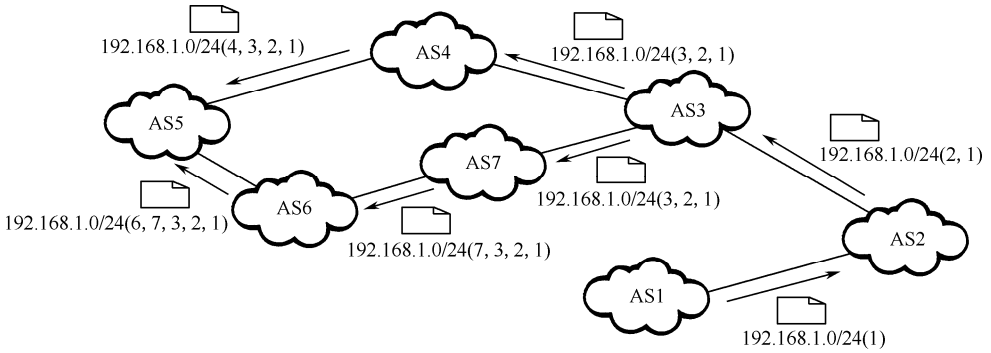


图 12.5　AS 路径列表

AS 列表并不记录 AS 内的路径信息，BGP 只关心 AS 间的路径，不关心 AS 内的路径信息。AS 内的路径信息由 IGP 提供。

根据 BGP 选路原则，具有最短 AS 列表的路由被认为是最优的路由。在如图 12.5 所示网络中，AS5 从 AS4 和 AS6 分别收到网络 192.168.1.0/24 路由通告的 AS 列表为“4,3,2,1”和“6,7,3,2,1”，显然，携带有“4,3,2,1”列表的路由是最优路由，因为它具有更短的 AS 列表。

第二，BGP 的路由环路检测。路由环路严重影响数据包在网络正常传递，必须采取一定措施加以避免。RIP 路由协议采取水平分割、毒性逆转、触发更新、时间抑制等措施避免路由环路发生。OSPF、IS-IS 路由协议是链路状态路由协议，其工作机制本身可以避免路由环路发生。那么 BGP 协议是如何避免路由环路发生的呢？BGP 协议通过检测收到的路由 AS 列表，如果 AS 列表中已经包含了自己的 AS 号，说明该路由已经通过了本地 AS 并被通告出去了。如果接收它将导致路由环路，所以将收到的带有本地 AS 号的路由丢弃。BGP 通过检测 AS 列表属性很容易地检测路由环路。

如图 12.6 所示的网络中，AS5 中的路由器从 AS4 收到了网络 192.168.1.0/24 的 BGP 路由，其 AS 列表为“4,3,2,1”。AS5 又将该路由通告给 AS6，并将自己的本地 AS 号添加到 AS 列表的起始位置，AS 列表为“5,4,3,2,1”。AS6 收到此 BGP 路由后又将该路由通告给 AS4，并将自己本地 AS 号添加到 AS 列表的起始位置，AS 列表为“6,5,4,3,2,1”。AS4 中的路由器从 AS6 中收到该 BGP 路由后，通过比较发现 AS 列表“6,5,4,3,2,1”有自己的 AS 号“4”。根据 BGP 的路由环路避免机制，AS4 的路由器拒绝接收此条 BGP 路由，将该路由丢弃，这样就避免了路由环路产生。

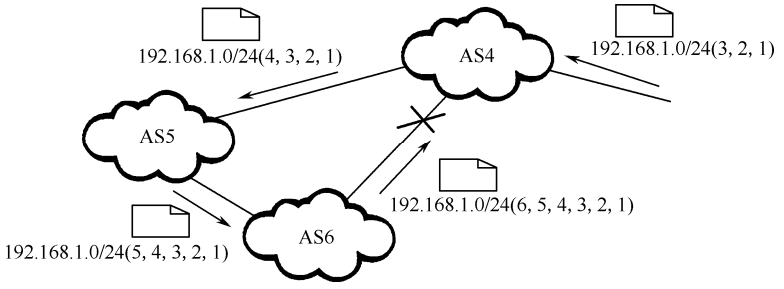


图 12.6　检测 BGP 路由环路

（4）BGP 的邻居关系。

通过前面的学习我们已经知道，RIP 路由协议使用 UDP 作为传输协议，接口号为 520。OSPF 路由协议使用 IP 协议作为传输协议，协议号为 89。那么 BGP 路由协议是使用什么协议来传输路由更新报文的呢？BGP 路由协议使用 TCP 作为传输协议，接口号为 179。由于 TCP 是面向连接的、可靠的

传输协议，这样可以简化 BGP 路由协议的设计。

在 BGP 术语中，将运行 BGP 协议的路由器称为 BGP 发言者（BGP Speaker）。在 Internet 中存在着成千上万的 BGP 发言者。每个发言者都可以与其他发言者建立邻居关系，如果属于一个自治系统（AS 号相同）的两个 BGP 发言者建立的邻居关系我们称为 IBGP（Internal BGP）关系；如果不属于一个自治系统（AS 号不相同）的两个 BGP 发言者建立的邻居关系我们称为 EBGP（External BGP）关系。

如图 12.7 所示的网络中，AS1 的发言者 R1 与 AS2 的发言者 R2、AS2 的发言者 R4 与 AS3 的发言者 R5 建立 EBGP 邻居关系，原因是它们链路两端的发言者属于不同的 AS。AS2 中的发言者 R2、R3、R4 之间建立 IBGP 邻居关系，同样 AS3 中的发言者 R5、R6 之间也建立 IBGP 邻居关系。

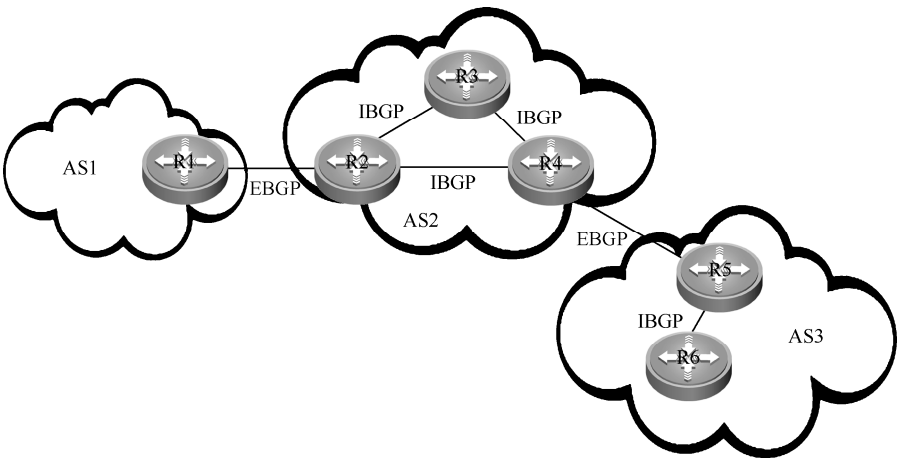


图 12.7　IBGP 与 EBGP

第一，EBGP 邻居关系。建立 EBGP 的目的是将路由信息从一个自治系统传递到另一个自治系统，也就是在自治系统之间交换路由选择信息。

默认情况下，BGP 规范中要求建立 EBGP 邻居关系的两个 BGP 发言者都必须是物理上直接连接的。但是我们也可以通过设置改变 EBGP 的这种默认行为，使其可以在非直接连接的两个 BGP 发言者之间建立 EBGP 邻居关系。当非直接连接的两个 BGP 发言者之间建立 EBGP 邻居关系时，我们称之为 EBGP 多跳（EBGP-Multihop）。

在图 12.8 所示的网络中，路由器 R1 和 R2 处于不同自治系统，并且它们之间通过直连的物理链路建立了 EBGP 邻居关系。路由器 R4 与 R5 也处于不同的自治系统，并且也通过物理链路直接连接，但是 R5 由于设备原因不支持 BGP 协议，所以无法与 R4 建立 EBGP 邻居关系。但是 AS3 希望与 AS2 建立邻居关系，接收 Internet 路由信息。在这种情况下，我们可以使用 EBGP 多跳机制，在路由器 R4 和 R6 之间建立 EBGP 多跳邻居关系。这样 AS3 就可以通过 R6 的 EBGP 关系接收 Internet 路由信息，然后再通过 BGP 路由重发布技术将 Internet 路由通告到路由器 R5。

在将路由器 R4 与 R6 建立 EBGP 多跳邻居关系之前必须实现 R4 与 R6 之间路由可达，我们可以采取配置静态路由或 IGP 的方式实现路由可达。IGP（RIP、OSPF 等）建立邻居关系必须在直连设备之间建立，而 BGP 不同，只要实现 TCP 连接，就可以建立邻居关系。

第二，IBGP 邻居关系。IBGP 的作用是用来在自治系统内部的 BGP 发言者之间传递 BGP 路由信息，使得自治系统内部的路由器也能获得到达自治系统外部的路径信息。

IBGP 与 EBGP 不同，IBGP 在建立邻居关系时不需要 BGP 发言者是直接连接的，只要两个 BGP 发言者之间路由可达即可。

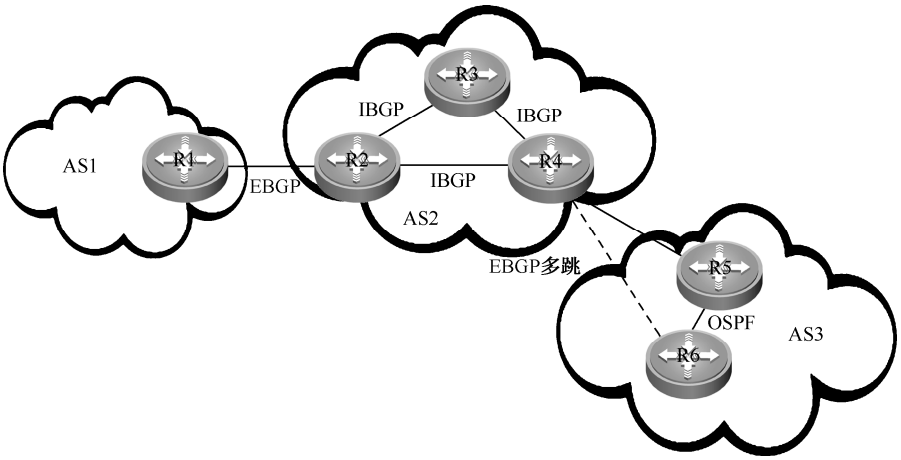


图 12.8　EBGP 多跳

① BGP 水平分割。前面介绍过 BGP 利用 AS 路径列表属性记录路由报文所走过 AS 的“路程”。在 AS 内部由于所有的路由器都拥有相同的 AS 号，路由报文所经过路由器的 AS 号都相同，所有无法使用 AS 列表来记录路由所走过的“路程”，如果不采取一定的措施将会产生路由环路现象。BGP 在 AS 内部的 IBGP 对等体之间传递路由信息时遵循这样一个规则：IBGP 发言者不能将从其他 IBGP 对等体接收到的路由信息再通告给另一个 IBGP 对等体，这种机制叫做“BGP 水平分割”。BGP 水平分割用来在 AS 内部防止 BGP 路由环路的产生。

在图 12.9 所示的网络中，在 AS2 的 R2、R3、R4 发言者建立了 IBGP 邻居关系。R2 从 AS1 的发言者 R1 中收到了一个 BGP 路由信息，然后将此路由信息通告给了本地自治系统的 IBGP 对等体 R3 和 R4。但是 R3 从对等体 R2 收到了这条路由信息后不会再将其通告给其他 IBGP 对等体 R4，这样避免了此路由信息在 R2、R3、R4 间形成路由环路。

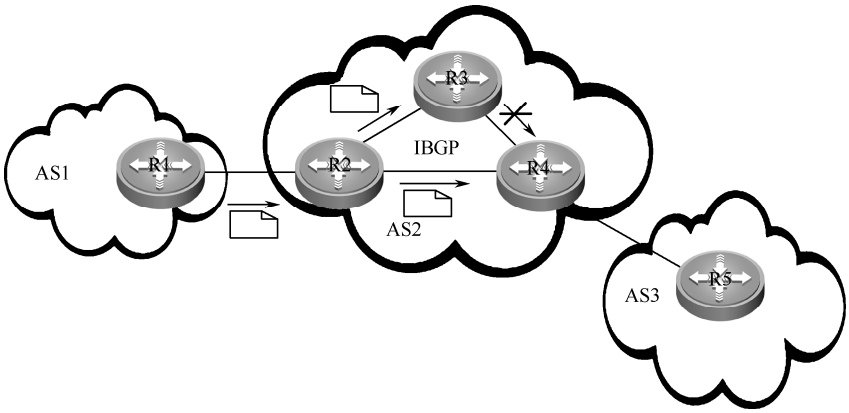


图 12.9　BGP 水平分割

从上述分析可以看出，由于路由信息的通告受到 BGP 水平分割的限制，为了确保在自治系统内的所有 BGP 发言者都能收到路由信息，必须在自治系统内部任何两个 BGP 发言者之间都建立邻居关系，也就是建立全网互联（Full Mesh）的邻居关系。通常情况下，在 AS 内部如何进行 IBGP 部署要根据具体情况而定。

② 末节 AS 的 IBGP 部署。如果要部署 IBGP 的网络是一个末节 AS 网络，即通过 BGP 连接到一

个或多个 ISP。可以只在边界路由器上启用 BGP 协议以接收 Internet 路由。然后可以在此边界路由器上采用路由重发布技术将 BGP 路由通告到 IGP 中，但是这样做会使得大量的 Internet 路由进入到 IGP 中，会消耗大量的系统资源，也可能造成路由器的系统瘫痪。在末节 AS 网络中比较常用的做法是在边界路由器上从 IGP 通告默认路由，既能让 IGP 中的路由器访问 Internet，又能减少对路由器系统资源的过渡消耗，如图 12.10 所示。

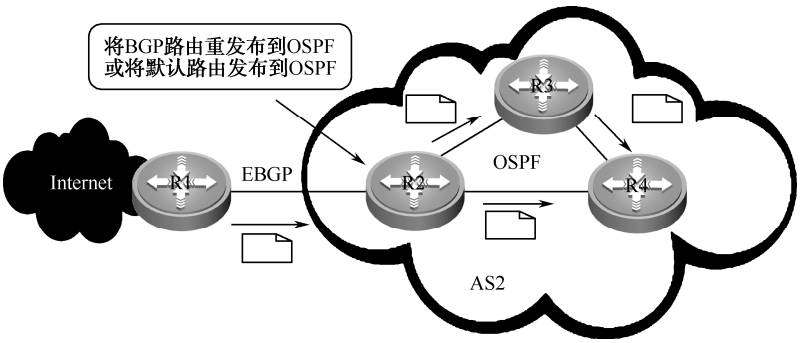


图 12.10　末节 AS 中的 IBGP 部署

③ 中转 AS 中的 IBGP 部署。如果要部署 IBGP 的网络是一个连接两个 AS 以上的多出口的中转网络，我们部署 IBGP 的方式有两种选择：一种是在整个 AS 内的路由器上都部署 BGP，建立全网互联网络；另一种是在边界路由器上部署 BGP，然后通过路由重发布技术将 Internet 路由发布到 IGP 中。

第一种部署 IBGP 方式要求 AS 内的所有路由器必须都支持 BGP 协议，并且路由器要有足够处理 Internet 大量路由信息的能力，如图 12.11 所示。

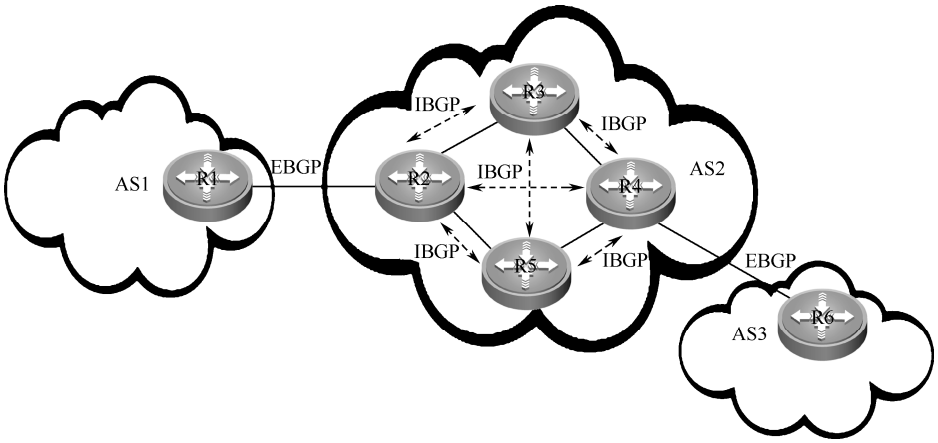


图 12.11　中转 AS 中的全网互联 IBGP 部署

在中转 AS2 中 BGP 发言者 R2、R3、R4、R5 之间建立了全网互联的 IBGP 邻居关系，这样 AS2 中的每台路由器都能接收到所有的外部路由信息。建立全网互联的 IBGP 邻居关系的前提是各个 BGP 发言者之间保证路由信息的可达性。

如果在上述 AS2 网络中，部分路由器不支持 BGP 协议，如 R3、R5 路由器，那么就不能实现 AS 全网互联的 IBGP 部署。此时可以采用第二种部署方式，即在边界路由器上部署 BGP，然后通过路由重发布通告外部路由信息。这种部署方式也要求 AS 内的所有路由器也要具备处理 Internet 大量路由信息的能力。

在如图 12.12 所示的网络中，在 AS2 的边界路由器 R2 和 R4 上都启用了 BGP 路由协议，以使其

能够从邻居 AS 中接收 Internet 路由信息。并且 R2 和 R4 也启用了 OSPF 路由协议。但是 R3 和 R5 路由器由于某种原因不支持 BGP 协议，只支持 OSPF 路由协议。为了使路由信息能够穿过本地 AS 并通告到其他 AS 中，R2 和 R4 建立了 IBGP 邻居关系。

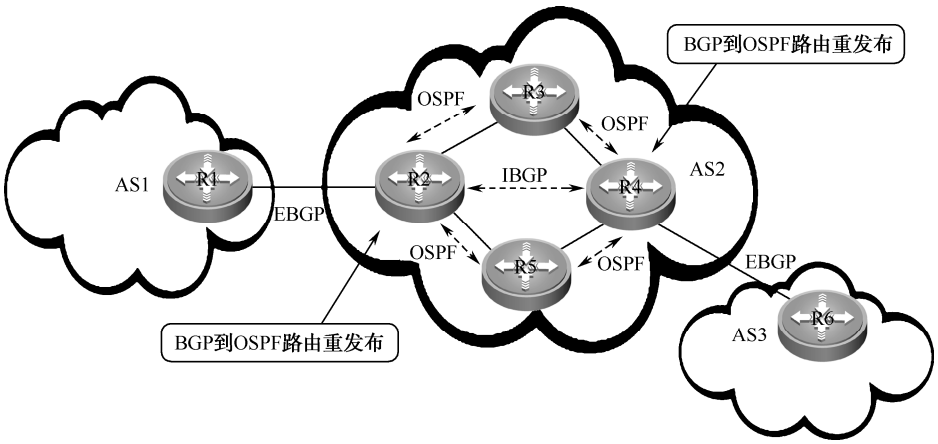


图 12.12　中转 AS 中的部分互联 IBGP 部署

由于 R3 和 R5 没有启用 BGP 协议，不能够直接获得外部路由信息，也就无法转发外部数据信息。因此，必须在边界路由器 R2 和 R4 上启动路由重发布，将外部路由信息通告到 OSPF 路由器中，这样一来路由器 R3、R5 能够获得外部路由信息，也就能顺利转发外部数据了。

（5）BGP 同步规则。

在上述描述中 AS2 作为中转 AS，在边界路由器 R2 和 R4 中同时启用 BGP 和 OSPF 路由协议，并且使用路由重发布将 AS 外部路由通告到 OSPF 中。这样做造成的后果是将会有大量的外部路由进入到 OSPF 中，如果 AS 内部路由器没有足够的处理能力将会 Down 机。如果在 AS2 中的所有路由器都启用 OSPF 协议，保证所有路由器之间的可达性。在 BGP 发言者 R2 和 R4 之间建立 IBGP 邻居关系，实现 BGP 路由信息的传递，是不是也能实现其他 AS 的路由信息经过本 AS 传递呢？

在图 12.13 所示的网络中，AS2 中的路由器 R2 和 R4 启用了 BGP 协议，同时 AS2 内的所有路由器都启用了 OSPF 协议，以便保证路由器之间的连通性。BGP 发言者 R2 和 R4 建立了 IBGP 邻居关系。路由器 R2 和 R4 没有实行路由重发布。

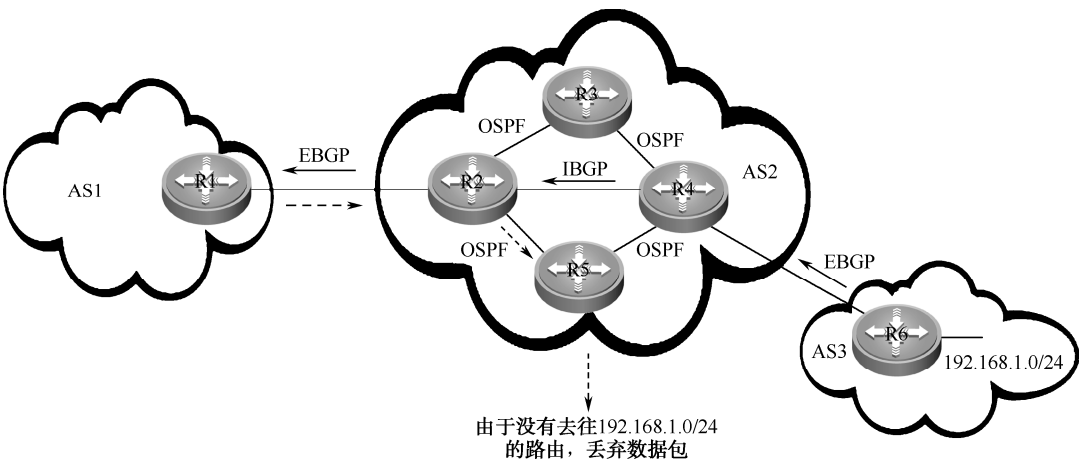


图 12.13　BGP 同步

AS3 中的 BGP 路由器将包含 192.168.1.0/24 网络的路由更新报文通告给 R4，R4 又将此更新报文通告给 IBGP 邻居 R2。R2 又通过 EBGP 连接将此路由更新报文通告给 AS1 中的路由器。现在 AS1 中的路由器获得了去往 AS3 中的 192.168.1.0/24 网络的路由信息。

但是当 AS1 中的发言者 R1 需要将数据包发送给 192.168.1.0/24 网络时，R1 通过查询路由表发现去往 192.168.1.0/24 网络的最佳路径是通过 AS2，于是将此数据包发送给 AS2 的 R2。R2 收到数据包后查询路由表发现去往 192.168.1.0/24 网络的最佳路径的下一跳是 R4，但是 R2 与 R4 没有直接连接，R2 查询路由表发现到达 R4 的最佳路径的下一跳是 R5，于是将数据包发送至 R5。R5 收到此数据包查询路由表，由于 R5 的路由表中不存在去往 192.168.1.0/24 网络的路由信息，于是将丢弃此数据包。R5 成为了网络中的路由黑洞，导致数据包不能被发送到目的地。

解决此问题的办法是让 R3、R5 学习到外部路由信息或者在该 AS 中的所有路由器都启用 BGP 协议。

为了防止 AS 网络中路由黑洞造成的数据丢失现象，BGP 定义了 BGP 同步规则，该规则要求：当 BGP 发言者通过 IBGP 接收到一条路由更新后，将检查此路由信息是否已经存在于 IP 路由表中，即是否通过 IGP 协议学习到了这条路由。如果 IP 路由表没有这条路由，那么它将不会再把这条路由通告给它的 EBGP 对等体，也就是不会通告给其他的 AS，除非已经学习到了这条路由。BGP 同步规则确保自治系统中的 BGP 与 IP 路由信息的一致性，防止丢失传递的 BGP 数据包。

默认情况下，路由器的同步功能是关闭的，需要时可以使用命令打开。当 AS 内的全部路由器都启用了 BGP 协议，并且所有 BGP 发言者之间都建立了 IBGP 邻居关系时可以关闭 BGP 的同步功能。另外作为末节 AS 网络也不需要 BGP 同步功能。

2. BGP 路径属性及配置命令

BGP 路由协议依据一系列路径属性判定路径的优劣，故也称为路径矢量路由协议。BGP 的路径属性可以分为：公认的与可选的，表示识别属性；强制的与自由决定的，表示携带属性；传递的与非传递的，表示传递属性。

其中，公认属性表示所有 BGP 发言者都必须能够识别它，这些属性在 BGP 更新报文中被传递给 BGP 对等体。

公认属性可以是强制的，也就是说，在所有的 BGP 更新报文中都要包含此属性；公认属性也是自由决定的，也就是说，在 BGP 更新报文中既可以出现，也可以不出现。

非公认的属性是可选的，可选属性可以是传递的或非传递的。

如果一个可选属性是传递的，即使 BGP 发言者不支持该属性，那么它也要接收包含该属性的路由，并把这个路由和属性传递给它的对等体。

如果一个属性是非传递的，不支持该属性的 BGP 发言者可以接收包含该属性的路由，但不会把这个属性传递给它的对等体。

（1）起源属性。

起源属性（ORIGIN）是一个公认强制属性，在 BGP 更新报文中的每条路由信息都将包含这个属性，而且接收它的 BGP 发言者也必须能够识别和处理这个属性。起源属性指明了这条路由信息的来源。当 BGP 发言者拥有多条到达相同目的地址的路由时，起源属性会作为一个决定路径优劣的因素。起源是指始发 BGP 发言者将这条路由通告到 BGP 中的方式。BGP 路由的起源属性可以有以下三种不同的选择。

第一，IGP 属性。IGP 起源属性表明此路由在始发 AS 内部自己产生的路由。当在路由器的 BGP 进程模式下使用 Network 命令将路由通告到 BGP 中时，这条路由的起源方式就是 IGP，在 BGP 路由

表中使用“i”表示。

第二，EGP 属性。EGP 起源属性表明此路由是在始发 AS 中由 EGP 协议学习到的路由。EGP 是 BGP 的前身路由协议，目前很少使用。在 BGP 路由表中使用“e”表示。

第三，Incomplete 属性。Incomplete 属性表明此路由是在始发 AS 中通过路由重发布方式获得的路由或源头未知的路由。如 BGP 发言者通过 OSPF 学习到一条 AS 内部路由，然后将其重发布到 BGP 中。在 BGP 路由表中使用“？”表示。

在如图 12.14 所示网络中，AS2 中的路由器 R3 拥有一条直连网络 192.168.1.0/24，R3 通过使用 Network 命令将直连路由通告到 BGP 中；路由器 R4 也拥有一条直连网络 192.168.2.0/24，R4 通过使用 redistribute 命令将直连路由重发布到 BGP 中。当 AS1 的路由器 R1 收到这两条路由后，看到 192.168.1.0/24 路由的起源属性为“i”，而 192.168.2.0/24 路由的起源属性为“？”。

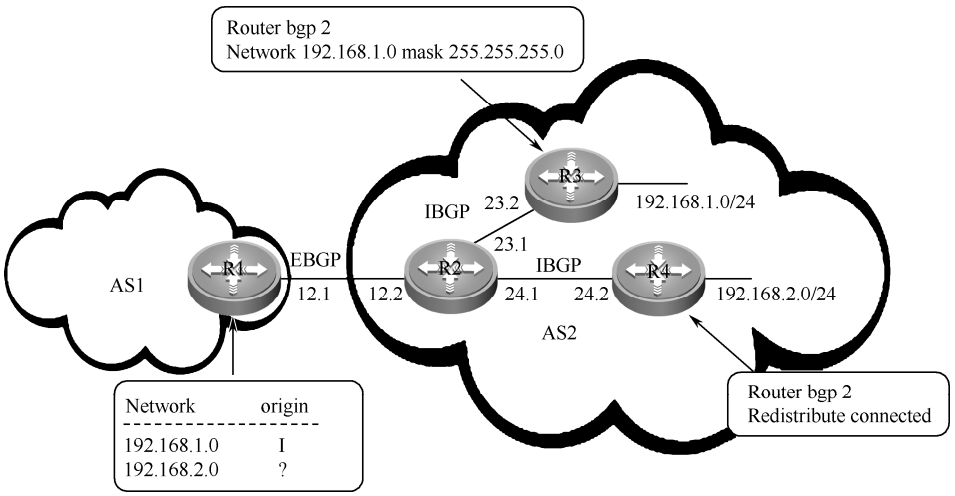


图 12.14　BGP 起源属性

针对上述三种路由起源方式，在路由选择中优先级顺序是 IGP>EGP>Incomplete，也就是说，当收到到达相同目的地的路由时，拥有 IGP 起源属性的路由将会优于 EGP 起源属性的路由，同理，拥有 EGP 起源属性的路由优于 Incomplete 属性的路由。

为了更深入地理解 BGP 属性，我们先学习几个基本配置命令，以便对 BGP 进行基本配置。

第 1 步：启用 BGP 路由协议。

在路由器上配置 BGP 路由协议时，首先需要启用 BGP 路由协议，并进入 BGP 路由协议配置模式。启用 BGP 路由协议的命令格式如下：

创建BGP路由进程命令

ruijie(config)#router bgp *as-number*

设备名称为ruijie
处于全局配置模式　　*自治系统号*

其中，*as-number* 为路由器所在的自治系统号，范围为 1～65535。默认情况下，关闭 BGP 路由协议。使用该命令的 no 选项关闭 BGP 协议，如 ruijie(config)#no router bgp 6500。

例如：创建自治系统号为 6500 的 BGP 路由进程。

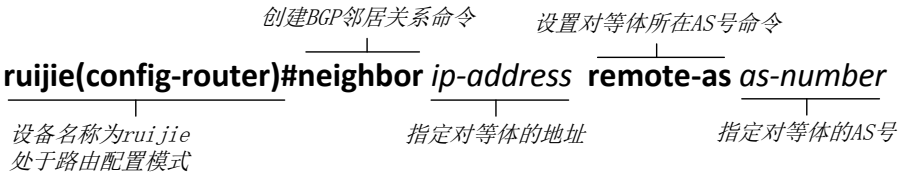
```
ruijie(config)#router bgp 65000
ruijie(config-router)#
```

例如：关闭自治系统号为 6500 的 BGP 路由进程。

```
ruijie(config)#no router bgp 65000
ruijie(config)#
```

第 2 步：配置 BGP 邻居。

通过前面的学习我们已经知道，BGP 是通过 TCP 建立与邻居的连接的，而 TCP 是点对点的单播可靠连接。所以不可能像 RIP、OSPF 等路由协议那样，可以自动建立邻居之间的连接。BGP 只能通过管理员手工地一个一个建立邻居连接关系。BGP 建立邻居关系的命令格式如下：



其中，*ip-address* 为指定对等体的 IP 地址；*as-number* 为指定对等体所在 AS 号。默认情况下，没有配置任何 BGP 对等体。使用该命令的 no 选项可删除配置的对等体，如 ruijie(config-router)#no neighbor 192.168.12.1 remote-as 2。

例如：创建与在 as 80 中 IP 地址为 10.0.0.1 的邻居连接。

```
ruijie(config)#router bgp 65000
ruijie(config-router)#neighbor 10.0.0.1 remote-as 80
ruijie(config-router)#
```

例如：删除上述建立的邻居连接。

```
ruijie(config)#router bgp 65000
ruijie(config-router)#no neighbor 10.0.0.1 remote-as 80
ruijie(config-router)#
```

第 3 步：向 BGP 协议注入路由信息。

虽然已经建立了邻居关系，但是 BGP 不会自动注入任何路由信息。可以通过 network 命令手工注入路由信息，也可以通过路由重发布技术注入路由信息。在这里仅介绍通过 network 命令注入路由方法。

在 BGP 路由进程模式下使用如下命令注入路由信息：

其中，*network-number* 为网络号；*mask* 为子网掩码。使用该命令的 no 选项删除设置的网络信息，如 ruijie(config-router)#no network 192.168.1.0 mask 255.255.255.0。

本命令允许将 IGP 的路由注入 BGP 的路由表中，公告的网络信息可以是直连路由、静态路由和动态路由。

例如：将直连路由 10.0.0.0 注入到 BGP 路由表中。

```
ruijie(config)#router bgp 65000
ruijie(config-router)#network 10.0.0.0 mask 255.255.0.0
```

需要特别注意的是，注入的路由信息必须是在 IGP 路由表中存在的路由，否则 BGP 不会注入到 BGP 路由表中，也不会通告给其他对等体。当 BGP 发言者从对等体收到 BGP 路由时，它将自动地把其通告给其他对等体。

例如，在路由器中进行了配置 ruijie(config-router)#network 172.16.1.0 mask 255.255.255.0，那么，BGP 将会在 IGP 路由表中查找 172.16.1.0/24 网络，如果路由表中不存在此网络，BGP 不会将其注入到 BGP 路由表中，也不会通告给其他对等体。如果在路由表中存在 172.16.0.0/16 或 172.16.1.128/25 网络，BGP 也不会通告 172.126.1.0/24 网络。

第 4 步：配置 BGP 自动汇总功能。

自动汇总功能将重发布到 BGP 的路由汇总到有类边界上，可以减小路由表的大小。使用如下命令进行路由自动汇总：

ruijie(config-router)#auto-summary

设备名称为ruijie
处于路由配置模式

默认情况下，BGP 不进行路由自动汇总。使用该命令的 no 选项将关闭该功能，如 ruijie(config-router)#no auto-summary。

例如：配置 BGP 路由自动汇总功能。

```
ruijie(config)#router bgp 65000
ruijie(config-router)#auto-summary
```

第 5 步：显示 BGP 路由表信息。

使用如下命令显示 BGP 路由表信息：

ruijie#show ip bgp

例如：显示 BGP 路由表信息。

```
ruijie#show ip bgp
Status codes: s suppressed, d damped, h history, * valid,
> best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Status      Network                          Next Hop Metric LocPrf     Path
------      ------------------               ---------------            --------
*>          211.21.21.0/24 110.110.110.10 0  1000                       200 300
*>          211.21.23.0/24 110.110.110.10 0  1000                       200 300
*>          211.21.25.0/24 110.110.110.10 0  1000                       300
*>          211.21.26.0/24 110.110.110.10 0  1000                       300
*>          211.21.27.0/24 110.110.110.10 0  1000                       200
```

上述显示的 BGP 路由表中，常有符号含义如下。

- “*”表示此路由是一条有效的合法路由。
- “>”表示此路由是到达目的地的最佳路由。
- “Next Hop”表示该路由的下一跳，也就是 BGP 下一跳属性。
- “Metric”表示 BGP 的 MED 属性值，“0”表示最优的 MED 值，也可以使用命令修改含义。
- “LocPrf”表示路由的本地优先级属性，默认的本地优先级为 100，默认值不显示。
- “Path”表示路由的 AS 路径属性。

第 6 步：BGP 复位命令。

由于 BGP 协议本身不具备自动建立邻居连接与路由信息更新机制，当路由策略或者 BGP 的配置发生改变，其邻居路由器也不能自动感知到配置的变化，必须通过命令重新实现邻居路由信息的更新。使用 BGP 的复位命令，可以在不关闭 BGP 会话连接的情况下，有效地实施新的路由策略。BGP 复位命令格式如下：

其中，“*”表示复位当前所有 BGP 会话。

例如：复位当前所有 BGP 会话。

```
ruijie#clear ip bgp *
```

提示：*当路由器的 BGP 配置修改时，为了将修改后的结果让邻居路由器能看到，必须在修改后使用 clear ip bgp *命令将 BGP 会话进行复位、更新，否则邻居路由器看到的效果还是修改前的内容。*

第 7 步：“起源属性”应用实例。

在如图 12.15 所示网络中，AS1 的路由器 R1 中 Loop0 通过 network 命令宣布，将 1.1.1.0/24 网络路由以 IGP 起源属性注入；Loop1 通过 redistribute 命令重发布，将 2.2.2.0/24 网络路由以 Incomplete 属性注入。具体配置代码如下。

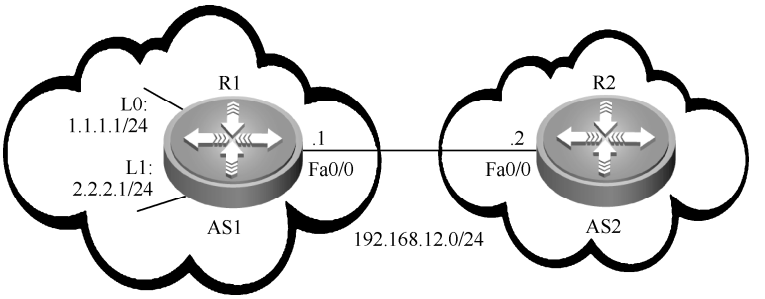


图 12.15　BGP 配置实例

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
R1(config)#interface Fastethernet 0/0
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 192.168.12.2 remote-as 2
R1(config-router)#network 1.1.1.0 mask 255.255.255.0
R1(config-router)#redistribute connected
R1(config-router)#end
R1#
```

路由器 R2 配置如下：

```
R2(config)#interface Fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
R2(config-router)#end
R2#
```

第 2 步：测试 BGP 效果。

显示路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 3, local router ID is 192.168.12.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*   1.1.1.0/24       0.0.0.0                  0                32768      ?
*>                   0.0.0.0                  0                32768      i
*> 2.2.2.0/24       0.0.0.0                  0                32768      ?
*> 192.168.12.0     0.0.0.0                  0                32768      ?
Total number of prefixes 3
```

从上述显示结果看出，通过重发布及 network 命令分别获得去往 1.1.1.0/24 的路由，但是 network 命令方式更优；通过重发布方式获得 2.2.2.0/24 及 192.168.12.0/24 网络的路由。由于上述路由都是本地路由，路径 Path 属性值为 0。

显示路由器 R1 的 IP 路由表如下：

```
R1#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       1.1.1.0/24 is directly connected, Loopback 0
C       1.1.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
C       192.168.12.0/24 is directly connected, fastethernet 0/0
C       192.168.12.1/32 is local host.
```

上述路由表中路由全部都是本地直连路由。

显示路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 192.168.24.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
```

```
                S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 1.1.1.0/24       192.168.12.1             0                    0    1 i
*> 2.2.2.0/24       192.168.12.1             0                    0    1 ?
*> 192.168.12.0     192.168.12.1             0                    0    1 ?
Total number of prefixes 3
```

从上述显示结果看出，从 AS1 中学习到 3 条 BGP 路由，其中 1 条来源为“i”，其余 2 条来源为“？”；路径 Path 属性为“1”，表示从 AS1 获得的路由。

显示路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [20/0] via 192.168.12.1, 01:30:45
B       2.2.2.0/24 [20/0] via 192.168.12.1, 00:37:18
C       192.168.12.0/24 is directly connected, fastethernet 0/0
C       192.168.12.2/32 is local host.
```

从上述显示结果看出，除直连路由外，还有 2 条从 BGP 学习的路由。

（2）AS 路径属性。

AS 路径属性（AS_PATH）记录了要达到目的地网络需要经过的自治系统，也是一个公认强制属性，并且必须包含在路由更新中，所有 BGP 发言者都必须识别并处理该属性。AS 路径属性（AS_PATH）包括两种类型：有序 AS 列表（AS_SEQUENCE）和无序 AS 列表（AS_SET）。

第一，有序 AS 列表（AS_SEQUENCE）描述。有序 AS 列表指出了该路径所经过的自治系统的先后顺序，即这条路由所走自治系统的“路程”，用 AS 号来表示，如“5,2,1”表示该路由从 AS 起源，经过了 AS1、AS2、AS5 三个自治系统。

BGP 发言者通过本 AS 的边界路由器将一条 BGP 路由通告给其他 AS 的 EBGP 对等体时，将在该路由 AS 路径属性的起始位置处加入本地 AS 号。其他 AS 收到此路由后再通告给另一个 AS 时，也将在该路由 AS 路径属性的起始位置处加入自己的 AS 号。

需要注意的是，只要通过 EBGP 通告 BGP 路由时才添加 AS 号，也就是说，只有将 BGP 路由从一个 AS 通告到另一个 AS 时才添加 AS 号。在一个 AS 内通过 IBGP 通告 BGP 路由时不添加 AS 号，在 AS 内部发送和接收 BGP 路由的 BGP 发言者拥有相同的 AS 号。

“有序 AS 列表”应用实例如下。

在如图 12.16 所示的网络中，AS3 的 R4 作为 192.168.1.0/24 路由的始发路由器，通过 network 命令将该路由发布到 BGP 中。R4 在将该路由发送给 R3 之前，会将自己的 AS 号“3”添加到 AS 路径列表中，然后通过 EBGP 发送给 R3。但是在发送之前的 AS 路径属性为空。当 R3 收到 R4 发送过来的这条 BGP 路由后，发现 AS 路径属性为“3”，表明此路由始发于 AS3。R3 会将此路由通过 IBGP 通告给 R2，但不会修改 AS 路径属性，故 R2 看到的 AS 路径属性仍为“3”。R2 收到此路由后通过 EBGP

通告给 AS1 的 R1，并且在发送给 R1 之前将自己的 AS 号“2”添加到 AS 路径属性的起始位置处，使得路径属性为“2,3”。R1 收到此路由后发现 AS 路径属性为“2,3”，表明如果要到达 192.168.1.0/24 网络需要经过 AS2 和 AS3。

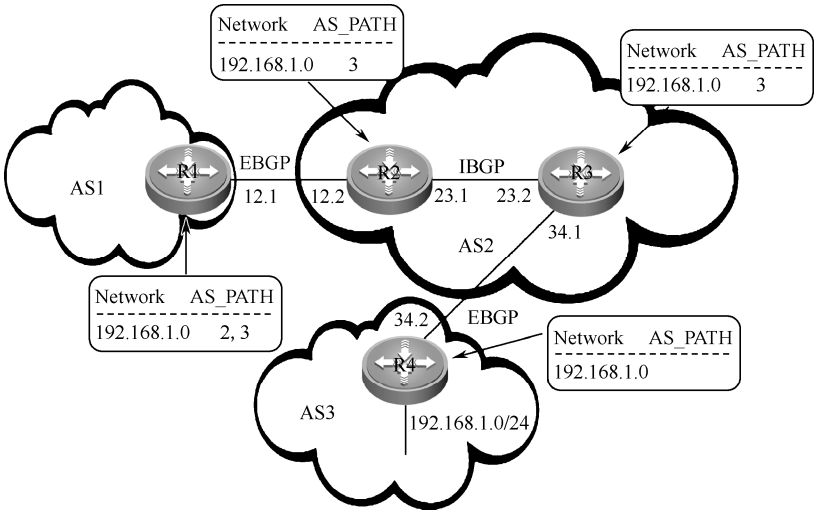


图 12.16　AS 路径属性

在路由选择中，BGP 路径属性中经过 AS 数量越少的路由越优，也就是说，在到达相同目的地的多条路径中，拥有最短 AS 路径列表的路径是最佳路径。

按照图 12.16 所示，我们对路由器进行 BGP 配置如下。

第 1 步：路由器 R1 配置代码如下。

```
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 192.168.12.2 remote-as 2
R1(config-router)#end
R1#
```

第 2 步：路由器 R2 配置代码如下。

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 192.168.23.1 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
R2(config-router)#neighbor 192.168.23.2 remote-as 2
R2(config-router)#end
R2#
```

第 3 步：路由器 R3 配置如下。

```
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 192.168.23.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface fastethernet 0/1
R3(config-if)#ip address 192.168.34.1 255.255.255.0
R3(config-if)#exit
R3(config)#router bgp 2
R3(config-router)#neighbor 192.168.23.1 remote-as 2
R3(config-router)#neighbor 192.168.34.2 remote-as 3
R3(config-router)#end
R3#
```

第 4 步：路由器 R4 配置如下。

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 192.168.34.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface loopback 0
R4(config-if)#ip address 192.168.1.1 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 3
R4(config-router)#neighbor 192.168.34.1 remote-as 2
R4(config-router)#network 192.168.1.0 mask 255.255.255.0
R4(config-router)#end
R4#
```

第 5 步：显示路由器 R4 的 BGP 路由表如下。

```
R4#show ip bgp
BGP table version is 2, local router ID is 192.168.34.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0/24   0.0.0.0                  0                 32768      i
Total number of prefixes 1
```

从上述显示结果看出，路由器 R4 学到起源属性为“i”的去往 192.168.1.0/24 网络的 BGP 路由信息。由于是本地路由，Path 属性为“0”。

显示路由器 R4 的 IP 路由表如下：

```
R4#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
```

```
C       192.168.1.0/24 is directly connected, Loopback 0
C       192.168.1.1/32 is local host.
C       192.168.34.0/24 is directly connected, FastEthernet 0/0
C       192.168.34.2/32 is local host.
```

从上述显示结果看出，路由器 R4 的 IP 路由表中全部是直连路由信息。

第 6 步：显示路由器 R3 的 BGP 路由表如下。

```
R3#show ip bgp
BGP table version is 2, local router ID is 192.168.34.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0/24   192.168.34.2             0                    0     3 i
Total number of prefixes 1
```

从上述显示结果看出，路由器 R3 从 AS3 的路由器 R4 学习到一条去往 192.168.1.0/24 网络的 BGP 路由。由于是从 AS3 学到的路由信息，path 属性值为“3”。

显示路由器 R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       192.168.1.0/24 [20/0] via 192.168.34.2, 00:09:57
C       192.168.23.0/24 is directly connected, FastEthernet 0/0
C       192.168.23.2/32 is local host.
C       192.168.34.0/24 is directly connected, FastEthernet 0/1
C       192.168.34.1/32 is local host.
```

从上述显示结果看出，路由器 R3 的路由表中学到一条 BGP 路由，管理距离值为默认值 20，下一跳地址为 192.168.34.2。

第 7 步：显示路由器 R2 的 BGP 路由表如下。

```
R2#show ip bgp
BGP table version is 23, local router ID is 192.168.23.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*>i192.168.1.0/24   192.168.34.2             0        100          0     3 i
Total number of prefixes 1
```

从上述显示结果看出，路由器 R2 学习到一条去往 192.168.1.0/24 网络的 BGP 路由信息，但是由于 R2 与 R3 之间建立的 IBGP 对等体，path 属性值仍然为“3”。

显示路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
```

从上述显示结果看出，路由器 R2 的 IP 路由表中并没有学习到去往 192.168.1.0/24 网络的 BGP 路由信息。为什么路由器 R2 已经通过 BGP 协议学习到去往 192.168.1.0/24 网络的路由信息却没有加入到 IP 路由表中呢？原因在于虽然 R2 通过 BGP 学习到去往 192.168.1.0/24 网络的 BGP 路由信息，但是在 IP 路由表中，没有去往下一跳地址（192.168.34.2）的路由，所以这条路由是无效路由，故不能加入到路由表中。解决的办法是让 R2 学到去往 192.168.34.2 的路由。

第 8 步：显示路由器 R1 的 BGP 路由表信息如下。

```
R1#show ip bgp
```

没有任何路由信息。AS1 中的路由器 R1 没有从 AS2 中学习到 BGP 路由信息，原因跟上面的叙述一样，R2 的去往 192.168.1.0/24 网络的 BGP 路由是无效的路由。无效的路由不能传递到 EBGP 对等体中，即 BGP 路由同步问题。

显示路由器 R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.1/32 is local host.
```

第 9 步：从解决上述无效 BGP 路由问题，我们在路由器 R2 和 R3 上分别启用 RIP 路由协议，使得路由器 R2 能够学到去往 192.168.1.0/24 网络路由的下一跳地址 192.168.34.2 的路径。

在路由器 R2 上增加的配置代码如下：

```
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#network 192.168.12.0
R2(config-router)#network 192.168.23.0
R2(config-router)#no auto-summary
```

在路由器 R3 上增加的配置代码如下：

```
R3(config)#router rip
R3(config-router)#version 2
R3(config-router)#network 192.168.23.0
R3(config-router)#network 192.168.34.0
R3(config-router)#no auto-summary
```

经过上述增加代码配置过程后，显示路由器 R1、R2 的路由表如下。

第 10 步：显示路由器 R1 的 BGP 路由表信息如下。

```
R1#show ip bgp
BGP table version is 2, local router ID is 192.168.12.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0/24   192.168.12.2             0                    0    2 3 i
Total number of prefixes 1
```

从上述显示结果看出，路由器 R1 学习到一条去往 192.168.1.0/24 网络的 BGP 路由信息，但是由于从 AS2 的 R2 中学到的，path 属性值仍然为“2 3”。下一跳地址为 AS 中的路由器 R2 接口地址 192.168.12.2。

显示路由器 R1 的 IP 路由表如下：

```
R1#show ip rou
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       192.168.1.0/24 [20/0] via 192.168.12.2, 00:01:48
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.1/32 is local host.
```

从上述显示结果看出，路由器 R1 的路由表中学到一条 BGP 路由，由于该路由是从 EBGP 对等体中学到的，故管理距离值为默认值 20，下一跳地址仍然为 192.168.12.2。

显示路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       192.168.1.0/24 [200/0] via 192.168.34.2, 00:01:34
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
```

```
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
R       192.168.34.0/24 [120/1] via 192.168.23.2, 00:02:28, FastEthernet 0/1
```

从上述显示结果看出，路由器 R2 的路由表中学到一条 BGP 路由，由于该路由是从 IBGP 对等体中学到的，故管理距离值为默认值 200，下一跳地址仍然为 192.168.34.2。

第二，无序 AS 列表（AS_SET）描述。AS 路径属性除是路由选择时的重要参数外，还是防止路由环路的必要措施。如果 BGP 发言者接收到来自 EBGP 对等体路由的 AS 路径列表中包含本地的自治系统号，那么它将丢弃该路由更新。但是当对多个详细的路由进行聚合后，会丢失一些 AS 路径的详细信息，结果使导致路由环路的概率大大增加。

在如图 12.17 所示的网络中，AS5 接收到 AS4 通告的具有聚合属性的路由 172.16.0.0/26（4）后，又将其 172.16.0.0/26（5,4）通告给了 AS3。AS3 中的 BGP 发言者收到这条路由并查看路由的 AS 路径列表，由于 AS 路径列表（5,4）中没有本地 AS 号（3），所以它将接受这条更新。这时如果 AS3 中的一个网络 172.16.2.16/28 出现故障，那么 AS3 中去往该网络的数据流都将使用 AS5 接收的聚合路由并发送给 AS5，然后 AS5 再将数据发送给 AS4，AS4 又使用 172.16.2.0/24 将数据重新发送回 AS3，这样就出现了路由环路。由于聚合器后面的详细路径信息没有包含在 AS 路径列表中，所以 AS3 不能检测出环路。

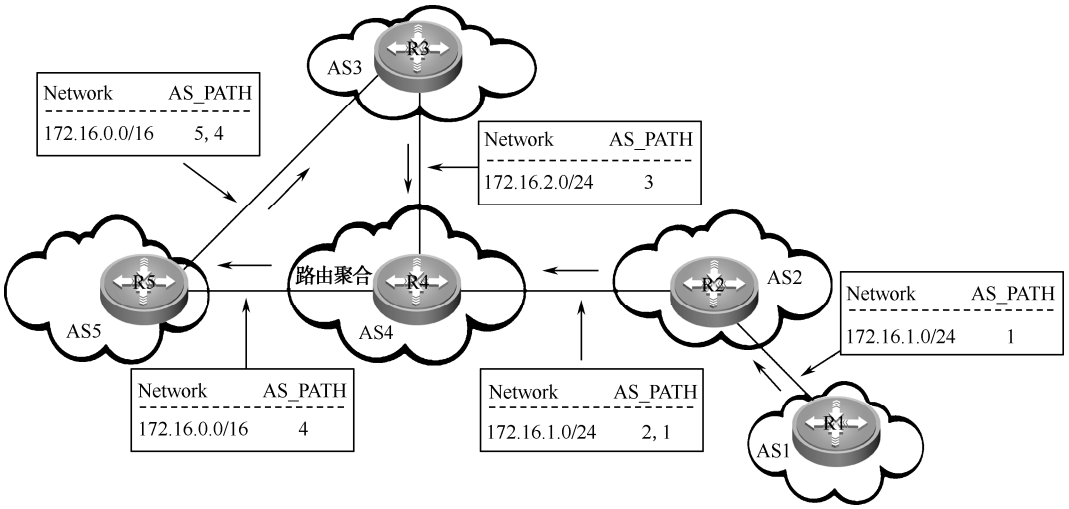


图 12.17　路径属性的丢失可能造成路由环路

对于上述产生路由环路问题，避免的方法是在 AS4 路由器上使用 AS_SET 属性，以便 AS3 及时发现本地 AS 号，拒绝接收 AS5 发回的聚合路由。

AS_SET 属性就是聚合器在聚会路由时产生聚合路由的同时插入源路由的详细路径信息。如图 12.18 所示，AS4 中的聚合器在对路由进行聚合时设置了 AS_SET 信息，AS_SET 信息中包含了源路由的详细路径信息（1,2,3）。在 BGP 路由表中，AS_PATH 中的 AS_SET 信息使用括号进行表示。同时聚合器也设置了 AS_SEQUENCE，即“4”。AS_SEQUENCE 也就是我们之前看到的 AS 路径列表，它记录了以聚合器所在自治系统为起点的 AS 路径列表。需要注意的是，AS_SET 是一个无序的列表，我们也可以看出，详细路由的路径信息处于不同的分支上，无法使用有序列表来描述路径信息。

当 AS5 收到这条聚合路由并把它通告给 AS3 时，由于 AS3 中的 BGP 发言者看到了在 AS_PATH

中 AS_SET 路径信息中包含自己的 AS 号，所以它将忽略这条更新，这样也就避免了产生路由环路。

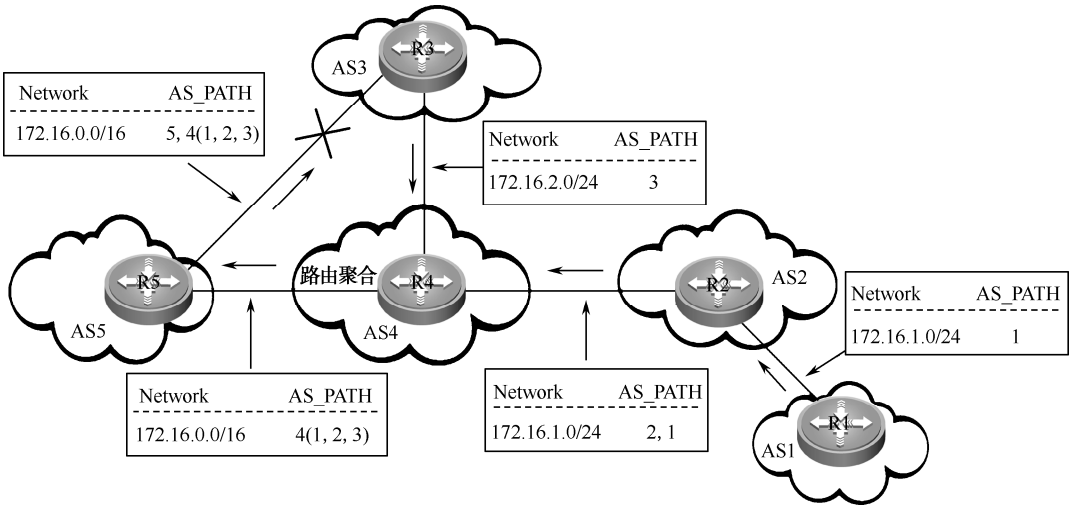


图 12.18　AS_SET 属性

BGP-4 支持 CIDR，所以允许创建聚合表项，以减小 BGP 路由表的大小。当然，只有当聚合范围内存在有效的路径时，才将 BGP 聚合表项添加到 BGP 路由表中。

在路由进程配置模式下，使用如下命令配置 BGP 路由聚合：

设置BGP路由聚合命令　　*指定聚合后的子网掩码*　　*只通告聚合后的路由*

ruijie(config-router)#aggregate-address *ip-address mask* **{as-set|summary-only}**

设备名称为ruijie
处于路由配置模式　　*指定聚合后的网络地址*　*包含详细路由的AS路径*

其中，*ip-address* 为聚合后的网络地址；*mask* 为聚合后的子网掩码；*as-set* 选项：当使用该选项时，聚合路由中包含详细路由的 as 路径信息，即 as-set。**summay-only** 选项：当使用该选项时，只通告聚合后的路由信息。默认情况下，既通告聚合后的路由，也通告详细的路由。

默认情况下，没有配置聚合功能。使用该命令的 no 选项，将关闭该功能，如 ruijie(config-router)# no aggregate-address 172.16.0.0 255.255.0.0 summary-only。

“无序 AS 列表”应用实例如下。

在如图 12.19 所示网络中，在 AS2 中的路由器 R2 配置 BGP 路由聚合，并且使用 summary-only 选项，只通告聚合后的路由信息。

第 1 步：进行不带任何选项的聚合命令 aggregate-address 配置。

路由器 R1 配置如下：

```
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 192.168.12.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 192.168.12.2 remote-as 1
R1(config-router)#end
R1#
```

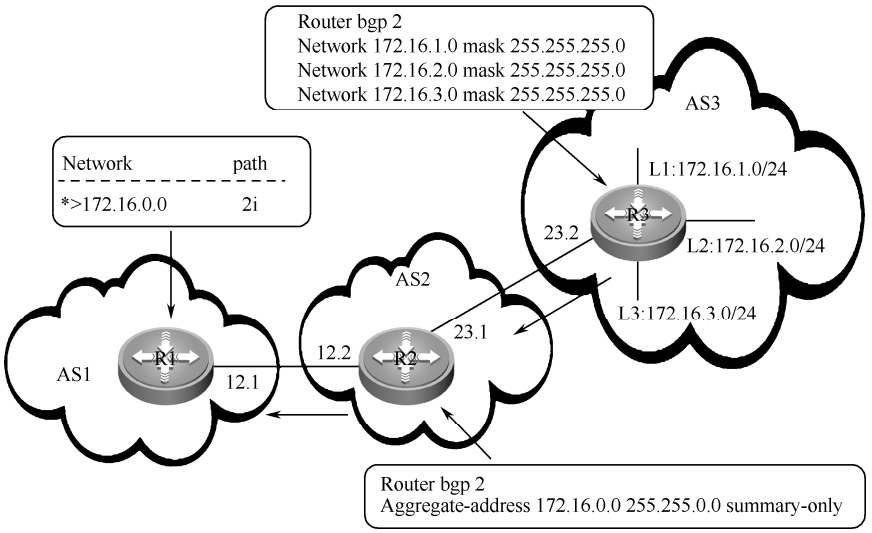


图 12.19　配置 BGP 路由聚合

路由器 R2 配置如下：

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 192.168.12.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 192.168.23.1 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
R2(config-router)#neighbor 192.168.23.2 remote-as 2
R2(config-router)#aggregate-address 172.16.0.0 255.255.0.0
R2(config-router)#end
R2#
```

路由器 R3 配置如下：

```
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 192.168.23.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface loopback 1
R3(config-if)#ip address 172.16.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#interface loopback 2
R3(config-if)#ip address 172.16.2.1 255.255.255.0
R3(config-if)#exit
R3(config)#interface loopback 3
R3(config-if)#ip address 172.16.3.1 255.255.255.0
R3(config-if)#exit
R3(config)#router bgp 3
R3(config-router)#neighbor 192.168.23.1 remote-as 2
```

```
R3(config-router)#network 172.16.1.0 mask 255.255.255.0
R3(config-router)#network 172.16.2.0 mask 255.255.255.0
R3(config-router)#network 172.16.3.0 mask 255.255.255.0
R3(config-router)#end
R3#
```

在路由器 R2 上使用 R2(config-router)#aggregate-address 172.16.0.0 255.255.0.0 命令进行路由聚合后，路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 16, local router ID is 192.168.12.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       192.168.12.2             0                  0    2 i
*> 172.16.1.0/24    192.168.12.2             0                  0    2 3 i
*> 172.16.2.0/24    192.168.12.2             0                  0    2 3 i
*> 172.16.3.0/24    192.168.12.2             0                  0    2 3 i
Total number of prefixes 4
```

从上述显示结果看出，默认情况下路由器 R1 从 R2 学习到聚合路由及详细路由信息，且汇总路由中 AS 列表从使用聚合命令的路由器 R2 开始。此情况下，路由聚合能减少路由数量的优点没有体现出来。显示 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [20/0] via 192.168.12.2, 00:04:09
B       172.16.1.0/24 [20/0] via 192.168.12.2, 00:05:22
B       172.16.2.0/24 [20/0] via 192.168.12.2, 00:05:22
B       172.16.3.0/24 [20/0] via 192.168.12.2, 00:05:22
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 192.168.23.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       0.0.0.0                                    32768      i
```

```
*> 172.16.1.0/24      192.168.23.2            0                      0    3 i
*> 172.16.2.0/24      192.168.23.2            0                      0    3 i
*> 172.16.3.0/24      192.168.23.2            0                      0    3 i
Total number of prefixes 4
```

以上显示结果表明，本地产生 1 条聚合路由及从路由器 R3 学习到 3 条详细 BGP 路由信息。

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [200/0] via 0.0.0.0, 00:03:17, Null 0
B       172.16.1.0/24 [20/0] via 192.168.23.2, 00:04:24
B       172.16.2.0/24 [20/0] via 192.168.23.2, 00:04:24
B       172.16.3.0/24 [20/0] via 192.168.23.2, 00:04:24
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
```

以上显示结果表明，4 条 BGP 路由皆加入到 IP 路由表中。

再看一下这种情况下，路由器 R3 的情况。路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 9, local router ID is 172.16.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf    Path
*> 172.16.0.0       192.168.23.1             0                2 i
*> 172.16.1.0/24    0.0.0.0                  0                  i
*> 172.16.2.0/24    0.0.0.0                  0                  i
*> 172.16.3.0/24    0.0.0.0                  0                  i
Total number of prefixes 4
```

从上述显示结果表明，路由器 R3 也从 R2 学习到聚合路由，从本地产生 3 条 BGP 路由。从 R2 学习到的聚合路由，包含本地路由信息，容易产生路由环路问题。R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
```

```
Gateway of last resort is no set
B       172.16.0.0/16 [20/0] via 192.168.23.1, 00:01:00
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/0
C       192.168.23.2/32 is local host.
```

上述使用没有任何选项的 aggregate-address 聚合路由命令，默认情况下同时发布聚合路由及详细路由信息，并没有发挥聚合路由能减少路由数量的优势。

第 2 步：使用携带 summary-only 选项的 aggregate-address 命令。

下面在路由器 R2 中将聚合命令中增加“summary-only”选项，看一下会出现什么效果呢？路由器 R2 中的 BGP 配置代码修改如下：

```
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
R2(config-router)#neighbor 192.168.23.2 remote-as 2
R2(config-router)#aggregate-address 172.16.0.0 255.255.0.0 summary-only
```

这样配置以后，路由器 R1、R2、R3 的路由信息如下：

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 17, local router ID is 192.168.12.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       192.168.12.2             0                    0    2 i
Total number of prefixes 1
```

从上述显示结果看出，路由器 R1 从 R2 只学习到聚合路由 172.16.0.0 信息，BGP 路由表中路由数量减少了；并且 AS 列表为“2”，表示从 AS2 发出的 BGP 路由信息。

路由器 R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [20/0] via 192.168.12.2, 00:07:10
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.1/32 is local host.
```

路由器 R1 的 IP 路由表也只有一条 BGP 聚合路由信息。

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 4, local router ID is 192.168.23.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       0.0.0.0                                      32768 i
s> 172.16.1.0/24    192.168.23.2             0                       0 3 i
s> 172.16.2.0/24    192.168.23.2             0                       0 3 i
s> 172.16.3.0/24    192.168.23.2             0                       0 3 i
Total number of prefixes 4
```

上述显示结果表明，3 条详细 BGP 路由信息被禁止发布；产生 1 条聚合路由信息，AS 列表为空。

路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [200/0] via 0.0.0.0, 00:00:15, Null 0
B       172.16.1.0/24 [20/0] via 192.168.23.2, 00:00:15
B       172.16.2.0/24 [20/0] via 192.168.23.2, 00:00:15
B       172.16.3.0/24 [20/0] via 192.168.23.2, 00:00:15
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
```

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 9, local router ID is 172.16.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Path
*> 172.16.0.0       192.168.23.1             0                2 i
*> 172.16.1.0/24    0.0.0.0                  0                  i
*> 172.16.2.0/24    0.0.0.0                  0                  i
*> 172.16.3.0/24    0.0.0.0                  0                  i
Total number of prefixes 4
```

上述显示结果表明，路由器 R3 也从 R2 学到 1 条 BGP 聚合路由信息，本地产生 3 条 BGP 路由信息，还是存在 BGP 路由环路问题。

路由器 R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [20/0] via 192.168.23.1, 00:07:52
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/0
C       192.168.23.2/32 is local host.
```

第 3 步：使用携带 as-set 及 summary-only 选项的 aggregate-address 命令。

为了解决 BGP 路由环路及路由数量问题，使用携带 as-set 及 summary-only 选项的 aggregate-address 命令。路由器 R2 的 BGP 配置代码修改如下：

```
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
R2(config-router)#neighbor 192.168.23.2 remote-as 2
R2(config-router)#aggregate-address 172.16.0.0 255.255.0.0 as-set summary-only
```

路由器 R2 修改完以上代码后，查看路由表有何变化呢？

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 20, local router ID is 192.168.12.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       192.168.12.2             0                    0    2 3 i
Total number of prefixes 1
```

上述显示结果中，AS 列表由“2”变为“2 3”，表示将详细路由的 AS 号“3”添加到 AS 列表中。

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [20/0] via 192.168.12.2, 00:06:05
C       192.168.12.0/24 is directly connected, FastEthernet 0/1
C       192.168.12.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 7, local router ID is 192.168.23.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 172.16.0.0       0.0.0.0                                    32768 3 i
s> 172.16.1.0/24    192.168.23.2             0                     0 3 i
s> 172.16.2.0/24    192.168.23.2             0                     0 3 i
s> 172.16.3.0/24    192.168.23.2             0                     0 3 i
Total number of prefixes 4
```

上述显示结果表明，本地产生 1 条聚合路由，并且 AS 列表属性包含详细路由 AS 号“3”。

路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       172.16.0.0/16 [200/0] via 0.0.0.0, 00:05:03, Null 0
B       172.16.1.0/24 [20/0] via 192.168.23.2, 00:01:03
B       172.16.2.0/24 [20/0] via 192.168.23.2, 00:01:03
B       172.16.3.0/24 [20/0] via 192.168.23.2, 00:01:03
C       192.168.12.0/24 is directly connected, FastEthernet 0/0
C       192.168.12.2/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/1
C       192.168.23.1/32 is local host.
```

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 10, local router ID is 172.16.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Path
```

```
*> 172.16.1.0/24     0.0.0.0          0                 i
*> 172.16.2.0/24     0.0.0.0          0                 i
*> 172.16.3.0/24     0.0.0.0          0                 i
Total number of prefixes 3
```

上述显示结果表明，路由器 R3 没有从 R2 学习到聚合路由信息，原因是聚合路由信息中包含本自治系统号“3”，拒绝接收这样的聚合路由信息，从而避免 BGP 路由环路问题。

路由器 R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       172.16.1.0/24 is directly connected, Loopback 1
C       172.16.1.1/32 is local host.
C       172.16.2.0/24 is directly connected, Loopback 2
C       172.16.2.1/32 is local host.
C       172.16.3.0/24 is directly connected, Loopback 3
C       172.16.3.1/32 is local host.
C       192.168.23.0/24 is directly connected, FastEthernet 0/0
C       192.168.23.2/32 is local host.
```

实践结果证明，使用同时选择 as-set 及 summary-only 选项的 BGP 聚合 aggregate-address 命令，既能减少路由数量又能避免路由环路问题。

（3）下一跳属性。

下一跳属性（NEXT_HOP）也是一个公认强制属性，在所有 BGP 路由中都携带该属性，并且所有 BGP 发言者都必须识别和处理该属性。下一跳属性描述了要到达目的网络需要将数据包发送到的下一个地址或路由器。

在 BGP 协议中，EBGP 路由和 IBGP 路由的下一跳规则是不同的，下面分别进行叙述。

第一，EBGP 下一跳描述。对于 EBGP 来讲，当 BGP 发言者将路由更新报文发送给 EBGP 邻居时，将把该路由的下一跳属性设置为发送更新报文的路由器物理接口 IP 地址或回环接口地址。这个特性类似 RIP 路由协议。

“EBGP 下一跳”应用实例如下。

在如图 12.20 所示的网络中，R1 和 R2 路由器建立了 EBGP 邻居关系，R1 通过 EBGP 将网络 192.168.1.0/24 的路由通告给了 R2，并且在 R1 发送该路由更新之前将该路由的 BGP 下一跳属性修改为本路由器接口 IP 地址 12.1.1.1。当 R2 收到此路由更新时，发现下一跳属性为 12.1.1.1，获知要到达网络 192.168.1.0/24，需要将数据包发往 12.1.1.1 地址。

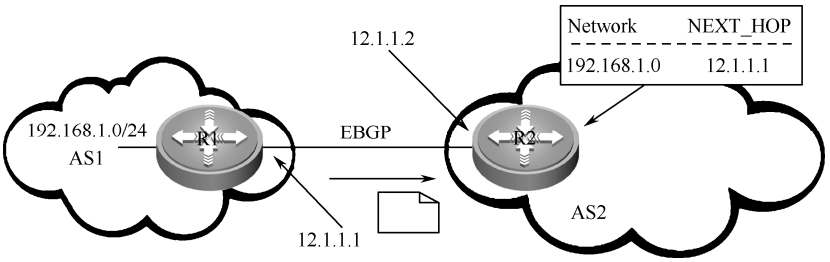


图 12.20 EBGP 下一跳属性

路由器 BGP 配置代码如下。

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 12.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface loopbakup 0
R1(config-if)#ip address 192.168.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 2
R1(config-router)#network 192.168.1.0 mask 255.255.255.0
R1(config-router)#end
R1#
```

路由器 R2 配置如下：

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#end
R2#
```

第 2 步：测试 BGP 效果。

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 1, local router ID is 192.168.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0      0.0.0.0                  0                 32768      i
Total number of prefixes 1
```

上述显示结果表明，路由器 R1 产生 1 条本地 BGP 路由信息。R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       12.1.1.0/24 is directly connected, FastEthernet 0/1
C       12.1.1.1/32 is local host.
C       192.168.1.0/24 is directly connected, Loopback 0
C       192.168.1.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 2, local router ID is 12.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0      12.1.1.1                 0                    0     1 i
Total number of prefixes 1
```

上述显示结果表明，AS2 中路由器 R2 从 AS1 中学习到 1 条 BGP 路由信息，下一跳地址为 12.1.1.1，即 EBGP 对等体的 IP 地址。R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
B       192.168.1.0/24 [20/0] via 12.1.1.1, 00:03:00
```

第二，IBGP 下一跳描述。IBGP 的下一跳的传输规则与 EBGP 不同。BGP 协议中规定，默认情况下，EBGP 通告的下一跳属性应该被传递到 IBGP 中。

“IBGP 下一跳”应用实例如下。

在图如 12.21 所示的网络中，路由器 R2 通过 EBGP 从 R1 处接收到 192.168.1.0/24 的路由更新报文，该路由的下一跳属性为 12.1.1.1，即 R1 发送更新路由的地址。R2 再将此路由通告给 IBGP 邻居 R3 时，依据 BGP 的规定，R2 不对此路由的下一跳属性进行修改，也就是说，R3 接收到这条路由的下一跳属性仍然为 12.1.1.1，而不是发送该路由更新报文的路由器接口地址 23.1.1.1。这是指 EBGP 通告的下一跳属性会被传递到 IBGP 中，也就是说，BGP 是一个以 AS 为跳数的路由协议。

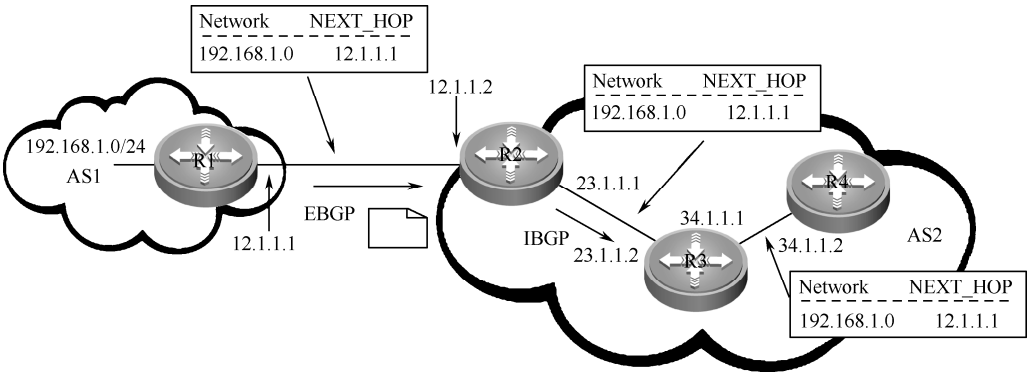


图 12.21 IBGP 下一跳

此时路由器 R3 知道如果要去往 192.168.1.0/24，应将数据发往 12.1.1.1，所以 R3 需要查找自己的路由表，寻址到达 12.1.1.1 的路由。如果 R3 存在去往 12.1.1.1 的路由，假设下一跳是路由器 R2，那么 R3 将会把去往 192.168.1.0/24 的数据发送给 R2。但实际情况可能并不是这样。如果 R3 的路由表中没有去往 192.168.1.0/24 的路由信息，那么 R3 会因为此条路由的下一跳不可达，而不会将此条路由加入到路由表中，这样去往 192.168.1.0/24 的数据包将被丢弃。

路由器 BGP 配置代码如下。

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
ruijie(config)#hostname R1
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 12.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface loopbackup 0
R1(config-if)#ip address 192.168.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 2
R1(config-router)#network 192.168.1.0 mask 255.255.255.0
R1(config-router)#end
R1#
```

路由器 R2 配置如下：

```
ruijie(config)#hostname R2
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 23.1.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 23.1.1.2 remote-as 2
```

```
R2(config-router)#neighbor 34.1.1.2 remote-as 2
R2(config-router)#exit
R2(config)#ip route 34.1.1.0 255.255.255.0 23.1.1.2
R2#
```

路由器 R3 配置如下：

```
ruijie(config)#hostname R3
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 23.1.1.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface fastethernet 0/1
R3(config-if)#ip address 34.1.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#router bgp 2
R3(config-router)#neighbor 23.1.1.1 remote-as 2
R3(config-router)#neighbor 34.1.1.2 remote-as 2
R3(config-router)#end
R3#
```

路由器 R4 配置如下：

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 34.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 2
R4(config-router)#neighbor 23.1.1.1 remote-as 2
R4(config-router)#neighbor 34.1.1.1 remote-as 2
R4(config-router)#exit
R4(config)#ip route 23.1.1.0 255.255.255.0 34.1.1.1
R4(config)#end
R4#
```

完成上述配置后，查询路由表，验证 BGP 效果。

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 1, local router ID is 192.168.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0      0.0.0.0                  0                 32768      i
Total number of prefixes 1
```

上述显示结果表明，路由器 R1 产生 1 条 BGP 路由信息。R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
```

```
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       12.1.1.0/24 is directly connected, FastEthernet 0/1
C       12.1.1.1/32 is local host.
C       192.168.1.0/24 is directly connected, Loopback 0
C       192.168.1.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 2, local router ID is 12.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 192.168.1.0      12.1.1.1                 0                     0     1 i
Total number of prefixes 1
```

上述显示结果表明，路由器 R2 从 R1 学习到 1 条 BGP 路由信息，下一跳地址为 12.1.1.1。R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
C       23.1.1.0/24 is directly connected, FastEthernet 0/1
C       23.1.1.1/32 is local host.
S       34.1.1.0/24 [1/0] via 23.1.1.2
B       192.168.1.0/24 [20/0] via 12.1.1.1, 00:49:00
```

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 34, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf    Path
* i192.168.1.0      12.1.1.1                 0        100    1 i
Total number of prefixes 1
```

上述显示结果表明，路由器 R3 从 R2 学习到 1 条 BGP 路由信息，但下一跳地址也为 12.1.1.1，并

非是直连路由 23.1.1.1。R3 的 IP 路由表如下：

```
R3#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       23.1.1.0/24 is directly connected, FastEthernet 0/0
C       23.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
```

路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 7, local router ID is 34.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric     LocPrf     Weight Path
* i192.168.1.0     12.1.1.1                 0        100          0    1 i

Total number of prefixes 1
```

上述显示结果表明，路由器 R4 从 R2 学习到 1 条 BGP 路由信息，但下一跳地址也为 12.1.1.1，并非是直连路由 34.1.1.1。R4 的 IP 路由表如下：

```
R4#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
S       23.1.1.0/24 [1/0] via 34.1.1.1
C       34.1.1.0/24 is directly connected, FastEthernet 0/0
C       34.1.1.2/32 is local host.
```

第 2 步：修改 BGP 路由下一跳地址。

为了解决这个问题，在边界路由器上使用 **neighbor** *ip-address* **next-hop-self** 命令改变 IBGP 下一跳通告规则，使得 BGP 发言者（R2）将 IBGP 下一跳属性更改为自己路由器 IP 地址（23.1.1.1），而不是从 EBGP 接收来的下一跳属性（12.1.1.1）。

在向指定 BGP 对等体分发路由时，使用该命令将路由信息的下一跳设置为本 BGP Speaker。

使用的 **neighbor** *ip-address* **next-hop-self** 命令格式如下：

设置下一跳地址命令

ruijie(config-router)#neighbor *peer-address* **next-hop-self**

设备名称为ruijie
处于路由配置模式

BGP对等体IP地址

其中，*peer-address* 为 BGP 对等体地址。默认情况下，没有打开该功能，使用该命令的 no 选项可以取消该设置，如 ruijie(config)# no neighbor 172.16.1.2 next-hop-self。

在如图 12.22 所示的网络中，路由器 R2 修改了 BGP 下一跳属性，当它将路由信息通告给路由器 R3 时，将原来的下一跳属性 12.1.1.1 修改为自己发送路由更新的地址 23.1.1.1。这样 R3 看到了此路由的下一跳地址为 23.1.1.1，即本地直连网络。由于 192.168.1.0/24 网络对于 R3 来讲是路由可达的，所以 R3 会将此路由加入到 IP 路由表中进行数据包的转发。

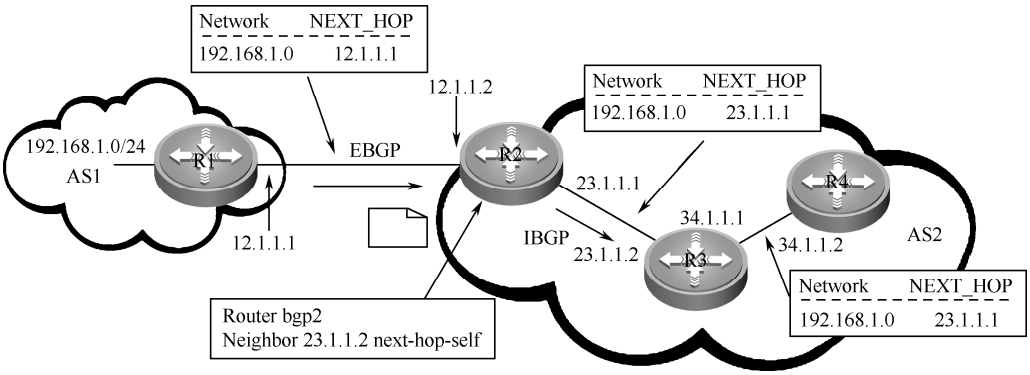


图 12.22　修改 IBGP 下一跳属性

为了实现上述功能，在 R2 路由器上添加了 neighbor 23.1.1.2 next-hop-self 命令；R1、R3、R4 路由器配置不变。

路由器 R2 配置修改如下：

```
ruijie(config)#hostname R2
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 23.1.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 23.1.1.2 remote-as 2
R2(config-router)#neighbor 34.1.1.2 remote-as 2
R2(config-router)#neighbor 23.1.1.2 next-hop-self
R2(config-router)#exit
R2(config)#ip route 34.1.1.0 255.255.255.0 23.1.1.2
R2(config)#end
R2#
```

完成上述 R2 的修改后，再次查看路由表。

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 40, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf  Path
*>i192.168.1.0      23.1.1.1                 0        100  1 i
Total number of prefixes 1
```

上述显示结果表明，路由器 R3 从 R2 学习到 1 条 BGP 路由信息，下一跳地址为 23.1.1.1，即 R3 的 IBGP 对等体地址。R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       23.1.1.0/24 is directly connected, FastEthernet 0/0
C       23.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
B       192.168.1.0/24 [200/0] via 23.1.1.1, 00:03:10
```

路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 20, local router ID is 34.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
* i192.168.1.0      12.1.1.1                 0        100          0  1 i
Total number of prefixes 1
```

上述显示结果表明，路由器 R4 从 R2 中学习到的 BGP 路由信息，下一跳地址仍然为 12.1.1.1。原因是在 R2 中只指定了 neighbor 23.1.1.2 next-hop-self。如果希望路由器 R4 从 R2 学习到的 BGP 路由的下一跳地址也是 12.1.1.1，在 R2 上需要再添加 neighbor 34.1.1.2 next-hop-self 即可。R4 的 IP 路由表如下：

```
R4#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
```

```
Gateway of last resort is no set
S       23.1.1.0/24 [1/0] via 34.1.1.1
C       34.1.1.0/24 is directly connected, FastEthernet 0/0
C       34.1.1.2/32 is local host.
```

如果 R3 再将这条路由通告给 EBGP 邻居时，将会把自己的路由器地址作为下一跳地址发送为 EBGP 邻居，如同 R1 发送给 R2 情况。

第 3 步：验证 AS 本地产生路由信息的下一跳地址。

刚才讨论的是 EBGP 通告的下一跳属性传播到 IBGP 中的规则，即从 EBGP 到 IBGP 的下一跳属性问题。如果在 AS 内的 BGP 发言者通过 IBGP 通告一条本地生成的路由时，下一跳地址将是生成该路由的 BGP 发言者的 IP 地址。

在如图 12.23 所示的网络中，10.1.1.0/24 路由由路由器 R4 始发生成，当 R4 将此路由通告给 IBGP 邻居 R3 时，会将下一跳属性设置为自己路由器接口地址 34.1.1.2。路由器 R3 收到这个路由后，将此路由通告给 IBGP 邻居 R2 时，下一跳属性仍然为 34.1.1.2。但是当路由器 R2 收到此路由后，将此路由通告给 EBGP 邻居 R1 时，将修改下一跳属性为自己的地址 12.1.1.2。

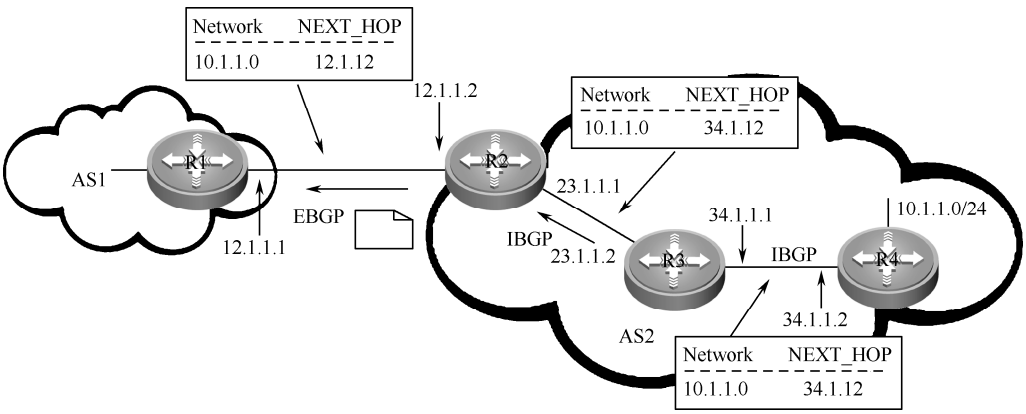


图 12.23　本地生成路由的下一跳属性

为了验证上述功能，修改路由器 R4 的配置如下：

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 34.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config-if)#interface Loopback 0
R4(config-if)#ip address 10.1.1.1 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 2
R4(config-router)#neighbor 23.1.1.1 remote-as 2
R4(config-router)#neighbor 34.1.1.1 remote-as 2
R4(config-router)#network 10.1.1.0 mask 255.255.255.0
R4(config-router)#exit
R4(config)#ip route 23.1.1.0 255.255.255.0 34.1.1.1
R4(config)#end
R4#
```

修改完 R4 的配置后，再次查看路由表信息。

路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 24, local router ID is 34.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 10.1.1.0/24      0.0.0.0                  0                    32768      i
*>i192.168.1.0      23.1.1.1                 0          100           0    1 i
Total number of prefixes 2
```

上述显示结果表明，路由器 R4 除从 R2 学到 1 条 BGP 路由信息外，还在本地产生 1 条 BGP 路由信息。路由器 R4 的 IP 路由表如下：

```
R4#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       10.1.1.0/24 is directly connected, Loopback 0
C       10.1.1.1/32 is local host.
S       23.1.1.0/24 [1/0] via 34.1.1.1
C       34.1.1.0/24 is directly connected, FastEthernet 0/0
C       34.1.1.2/32 is local host.
B       192.168.1.0/24 [200/0] via 23.1.1.1, 00:08:35
```

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 41, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf    Path
*>i10.1.1.0/24      34.1.1.2                 0          100      i
*>i192.168.1.0      23.1.1.1                 0          100    1 i
Total number of prefixes 2
```

上述显示结果表明，路由器 R3 从 R4 学习到 1 条 BGP 路由信息，下一跳地址为 34.1.1.2。路由器 R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
          i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
          ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       10.1.1.0/24 [200/0] via 34.1.1.2, 00:01:29
C       23.1.1.0/24 is directly connected, FastEthernet 0/0
C       23.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
B       192.168.1.0/24 [200/0] via 23.1.1.1, 00:16:56
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 12.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*>i10.1.1.0/24      34.1.1.2                 0        100          0      i
*> 192.168.1.0      12.1.1.1                 0                     0      1 i
Total number of prefixes 2
```

上述显示结果表明，路由器 R2 从 R4 学习到 1 条 BGP 路由信息，下一跳地址为 34.1.1.2。R2 的 IP 路由表如下：

```
R2#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       10.1.1.0/24 [200/0] via 34.1.1.2, 00:06:41
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
C       23.1.1.0/24 is directly connected, FastEthernet 0/1
C       23.1.1.1/32 is local host.
S       34.1.1.0/24 [1/0] via 23.1.1.2
B       192.168.1.0/24 [20/0] via 12.1.1.1, 01:11:10
```

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 2, local router ID is 192.168.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 10.1.1.0/24      12.1.1.2                 0                     0      2 i
```

```
*> 192.168.1.0          0.0.0.0                  0                      32768    i
Total number of prefixes 2
```

上述显示结果表明，路由器 R1 从 R2 学习到 1 条 BGP 路由信息，下一跳地址为 12.1.1.2。R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       10.1.1.0/24 [20/0] via 12.1.1.2, 00:07:13
C       12.1.1.0/24 is directly connected, FastEthernet 0/1
C       12.1.1.1/32 is local host.
C       192.168.1.0/24 is directly connected, Loopback 0
C       192.168.1.1/32 is local host.
```

（4）本地优先级属性。

本地优先级属性（LOCAL_PREF）是一个公认自由决定属性，也就是说，每个 BGP 发言者都必须能够识别该属性，但是 BGP 更新报文中可以不携带该属性。

本地优先级属性只会在 AS 内部传递，也就是说，本地优先级属性在 IBGP 对等体之间传递，它不会被通告给 EBGP 对等体，这也就是为什么称为“本地”所在。

本地优先级属性也是一个 BGP 用来进行路径决策的属性，优先级越高（数值越大）的路径被选为最佳路径的可能性越大。如果 BGP 发言者收到了多条到达相同目的地的路径，它将比较这些路径的本地优先级，选择本地优先级最高的路径作为最佳路径。BGP 发言者将接收到的外部路由发送给 IBGP 对等体时添加本地优先级属性。

第一，设置默认本地优先级描述。

默认的本地优先级为 100。默认的本地优先级可以通过如下命令进行修改：

ruijie(config-router)#bgp default local-preference *value*

设备名称为ruijie
处于路由配置模式

本地优先级值

其中，*value* 为设置的默认本地优先级值，范围为 0～4294967295。用该命令 no 选项来将其恢复默认值，如 ruijie(config-router)# no bgp default local-preference。

“设置默认本地优先级”应用实例如下。

本地优先级属性用于指导离开 AS 的数据包如何选择本地外出路径。在如图 12.24 所示的网络中，AS4 通告网络 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24。AS1 分别从 AS2 和 AS3 中收到这条路由更新报文。R3 收到此路由更新报文后通过 IBGP 发送给 R1，并且使用默认的本地优先级 100；R2 收到此路由更新报文后也通过 IBGP 发送给 R1，但是修改本地优先级为 150。R1 分别从 R2 和 R3 收到去往网络 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 的两条路由信息，根据 BGP 选路规则，BGP 发言者将选择本地优先级大的路径作为最佳路径，所以 R1 选择通过 R2 去往 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 的路径作为最佳路径。以后发往 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 的数据都将发往 R2。

本地优先级属性只具有本地意义，只会在本地 AS 内进行通告，不会将本地优先级属性通告给其他 AS。如果 R1 再将去往 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 的路由（本地优先级为 200）通告给其他 AS 的 EBGP 邻居时将不会携带本地优先级属性。

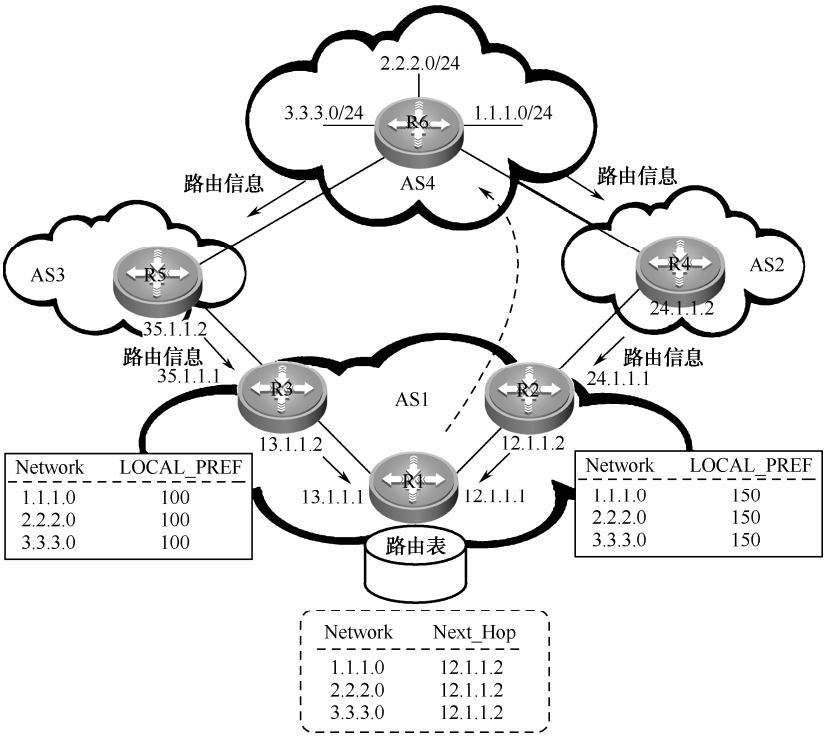


图 12.24　设置 BGP 默认本地优先级

为了方便实验，我们将如图 12.24 所示网络简化为如图 12.25 所示网络，但原理是一样的。

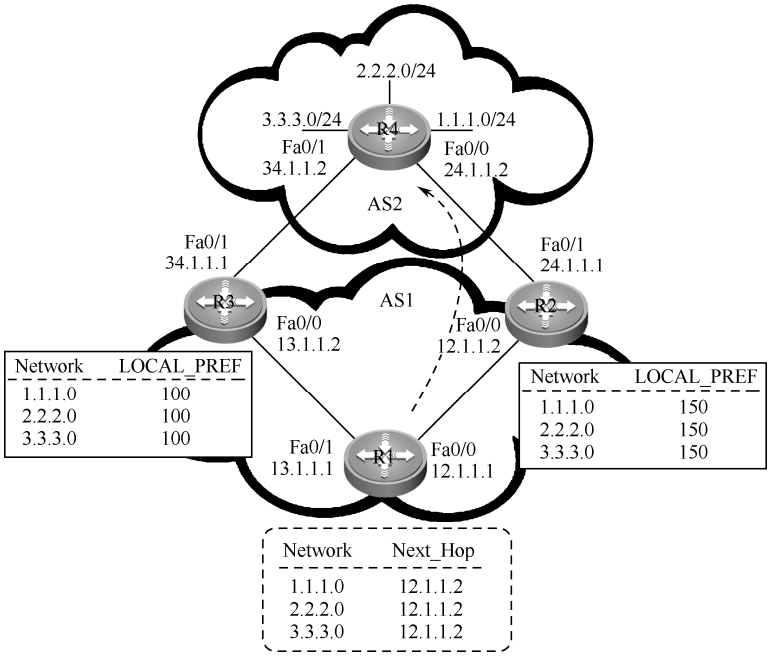


图 12.25　设置 BGP 默认本地优先级（简化）

针对如图 12.25 所示网络，配置相关的路由器。

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 12.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 13.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 1
R1(config-router)#neighbor 13.1.1.2 remote-as 1
R1(config-router)#exit
R1(config)#ip route 24.1.1.0 255.255.255.0 12.1.1.2
R1(config)#ip route 34.1.1.0 255.255.255.0 13.1.1.2
R1(config)#end
R1#
```

路由器 R2 配置如下：

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 24.1.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#router bgp 1
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 24.1.1.2 remote-as 2
R2(config-router)#bgp default local-preference 150
R2(config-router)#end
R2#
```

路由器 R3 配置如下：

```
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 13.1.1.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface fastethernet 0/1
R3(config-if)#ip address 34.1.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#router bgp 1
R3(config-router)#neighbor 13.1.1.1 remote-as 1
R3(config-router)#neighbor 34.1.1.2 remote-as 2
R3(config-router)#end
R3#
```

路由器 R4 配置如下：

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 24.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface fastethernet 0/1
R4(config-if)#ip address 34.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 0
R4(config-if)#ip address 1.1.1.1 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 1
R4(config-if)#ip address 2.2.2.1 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 2
R4(config-if)#ip address 3.3.3.1 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 2
R4(config-router)#neighbor 24.1.1.1 remote-as 1
R4(config-router)#neighbor 34.1.1.1 remote-as 1
R4(config-router)#network 1.1.1.0 mask 255.255.255.0
R4(config-router)#network 2.2.2.0 mask 255.255.255.0
R4(config-router)#network 3.3.3.0 mask 255.255.255.0
R4(config-router)#end
R4#
```

第 2 步：测试结果。

显示路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 3, local router ID is 3.3.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 1.1.1.0/24       0.0.0.0                  0                 32768      i
*> 2.2.2.0/24       0.0.0.0                  0                 32768      i
*> 3.3.3.0/24       0.0.0.0                  0                 32768      i
Total number of prefixes 3
```

上述显示结果表明，路由器 R4 产生了 3 条 BGP 路由信息。R4 的 IP 路由表如下：

```
R4#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
```

```
Gateway of last resort is no set
C       1.1.1.0/24 is directly connected, Loopback 0
C       1.1.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
C       3.3.3.0/24 is directly connected, Loopback 2
C       3.3.3.1/32 is local host.
C       24.1.1.0/24 is directly connected, FastEthernet 0/0
C       24.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.2/32 is local host.
```

显示路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 3, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf    Path
*> 1.1.1.0/24       34.1.1.2                 0                2 i
*> 2.2.2.0/24       34.1.1.2                 0                2 i
*> 3.3.3.0/24       34.1.1.2                 0                2 i
Total number of prefixes 3
```

上述显示结果表明，路由器 R3 从 R4 学习到 3 条 BGP 路由信息。R3 的 IP 路由表如下：

```
R3#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [20/0] via 34.1.1.2, 00:00:16
B       2.2.2.0/24 [20/0] via 34.1.1.2, 00:00:16
B       3.3.3.0/24 [20/0] via 34.1.1.2, 00:00:16
C       13.1.1.0/24 is directly connected, FastEthernet 0/0
C       13.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
```

显示路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 24.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 1.1.1.0/24       24.1.1.2                 0                    0    2 i
*> 2.2.2.0/24       24.1.1.2                 0                    0    2 i
*> 3.3.3.0/24       24.1.1.2                 0                    0    2 i
Total number of prefixes 3
```

上述显示结果表明，路由器 R2 从 R4 也学习到 3 条 BGP 路由信息。R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [20/0] via 24.1.1.2, 00:02:17
B       2.2.2.0/24 [20/0] via 24.1.1.2, 00:02:17
B       3.3.3.0/24 [20/0] via 24.1.1.2, 00:02:17
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
C       24.1.1.0/24 is directly connected, FastEthernet 0/1
C       24.1.1.1/32 is local host.
```

显示路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 3, local router ID is 13.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
  Network          Next Hop            Metric     LocPrf     Weight Path
* i1.1.1.0/24      34.1.1.2                 0        100          0    2 i
*>i                24.1.1.2                 0        150          0    2 i
* i2.2.2.0/24      34.1.1.2                 0        100          0    2 i
*>i                24.1.1.2                 0        150          0    2 i
* i3.3.3.0/24      34.1.1.2                 0        100          0    2 i
*>i                24.1.1.2                 0        150          0    2 i
Total number of prefixes 3
```

上述显示结果表明，路由器 R1 分别从 R2、R3 学习到 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 网络的 BGP 路由信息，但从 R2 学到的路由本地优先级为 150，从 R3 学到的路由本地优先级为默认值 100。R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
```

```
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [200/0] via 24.1.1.2, 00:00:23
B       2.2.2.0/24 [200/0] via 24.1.1.2, 00:00:23
B       3.3.3.0/24 [200/0] via 24.1.1.2, 00:00:23
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.1/32 is local host.
C       13.1.1.0/24 is directly connected, FastEthernet 0/1
C       13.1.1.1/32 is local host.
S       24.1.1.0/24 [1/0] via 12.1.1.2
S       34.1.1.0/24 [1/0] via 13.1.1.2
```

上述显示结果表明，路由器 R1 将优先级为 150 的 BGP 路由加入到 IP 路由表中。

第二，设置基于策略路由的 BGP 本地优先级描述。

由于上述 bgp default local-preference 命令不能针对不同的路由有选择地修改本地优先级，而是对从 EBGP 对等体收到的所有 BGP 路由的优先级都修改为 150，致使所有去往 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 的数据流量都从 R2 发往外部自治系统，R2 的负载过大，R1 的利用率过低。为了避免这种现象，必须使用路由策略设置，有选择地设置本地优先级。

下面介绍如何使用策略路由技术差异化配置本地优先级命令，达到 R2 和 R3 路由器流量均衡。

第 1 步：定义路由图。

要定义路由图，并进入路由图配置模式，请使用全局配置命令 route-map。

定义的路由图名称　　*路由图策略语句序号*

ruijie(config)#route-map *route-map-name* **[permit | deny]** *[sequence-number]*

设备名称为ruijie 处于全局配置模式　　*指定符合match条件，是否执行SET命令*

其中，*route-map-name* 为定义的路由图名称。

permit：如果定义了 permit 关键字，又符合 match 定义的匹配规则，则 set 命令将对数据包转发进行控制，并退出路由图的操作。如果定义了 permit 关键字，而不符合 match 定义的匹配规则，则将进入第二个路由图策略进行操作，直到最终执行了 set 命令

deny：如果定义了 deny 关键字，又符合 match 定义的匹配规则，则不会执行任何操作，该路由图策略不允许进行策略路由，而且退出路由图操作。如果定义了 deny 关键字，而不符合 match 定义的匹配规则，则将进入下一个路由图策略进行操作，直到最终执行了 set 命令。

sequence-number：路由图策略对应的序号。低序号的策略优先得到使用，因此需要注意序号的设置。

默认情况下，没有配置路由图。使用该命令的 no 形式删除指定路由图的定义，如 ruijie(config)# no route-map redrip permit 10。

例如：创建路由图 SET-PREF，对于从 EBGP 接收的 2.0.0.0 网络的路由设置本地优先级为 200，其他路由设置为默认值。

```
ruijie(config)#route-map SET_PREF permit 10
ruijie(config-route-map)#match ip address 1
```

```
ruijie(config-route-map)#set local-preference 200
ruijie(config-route-map)#exit
ruijie(config)#route-map SET_PREF permit 20
ruijie(config)#access-list 1 permit 2.0.0.0 0.255.255.255
```

第 2 步：设置策略条件。

使用如下命令设置策略路由的条件：

ruijie(config-route-map)#match ip address *access-list-number*

设备名称为ruijie
处于路由图配置模式

访问列表编号

其中，*access-list-number* 为访问列表编号。默认情况下，没有设置任何条件，使用该命令的 no 形式删除已有配置，如 ruijie(config-route-map)#no match ip address 10。

例如：对去往网络 200.168.23.0 中的路由进行设置。

```
ruijie(config)#access-list 10 permit 200.168.23.0 0.0.0.255
ruijie(config)#route-map redrip permit 10
ruijie(config-route-map)#match ip address 10
```

第 3 步：设置本地优先级。

要对匹配 match 规则的路由设置 LOCAL_PREFERENCE 值，使用路由图配置命令 set local-preference。该命令格式如下：

ruijie(config-route-map)#set local-preference *number*

设备名称为ruijie
处于路由图配置模式

本地优先级值

其中，*number* 为本地优先级数值，范围为 0～4294967295。

第 4 步：发布基于策略的 BGP 路由。

针对不同的对等体可以使用不同的 BGP 策略，命令格式如下：

对等体IP地址　　*数据流方向*

ruijie(config-router)#neighbor *peer-address* **route-map** *map-tag* {*in* / *out*}

设备名称为ruijie
处于路由配置模式

路由策略名称

其中，*peer-address* 为对等体的 IP 地址；*map-tag* 为路由策略名称；*in* 表示对进来的路由应用策略；*out* 表示对出去的路由应用策略。默认情况下，没有配置，使用该命令的 no 选项可以关闭该功能，如 ruijie(config-router)#no neighbor 192.168.12.1 route-map redrip in。

“设置基于策略路由的 BGP 本地优先级”应用实例如下。

在如图 12.26 所示网络中，AS1 网络希望去往 1.1.1.0/24 目标网络的数据流从 R2 路由器发送到 AS2，而去往 2.2.2.0/24、3.3.3.0/24 目标网络的数据流经 R3 发送到 AS3 中。为此，在 R2 路由器上，将去往 1.1.1.0/24 网络的路由本地优先级设置为 200，其他路由优先级为默认值 100；在 R3 路由器上将去往 2.2.2.0/24、3.3.3.0/8 网络的路由本地优先级设置为 200，其他路由优先级为默认值 100。

为了方便实验，我们将如图 12.26 所示网络简化为如图 12.27 所示网络，但原理是一样的。

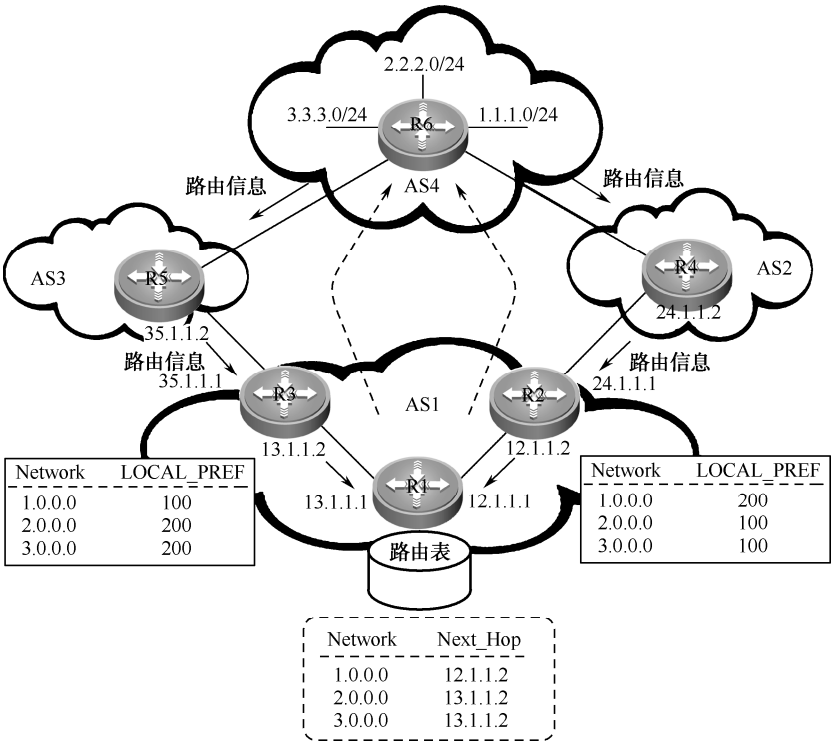


图 12.26　基于路由策略的本地优先级

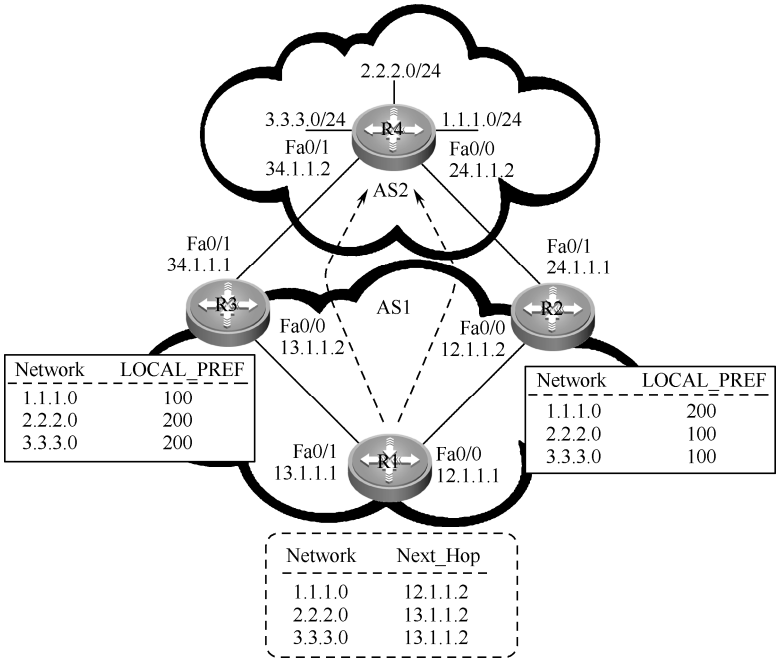


图 12.27　基于路由策略的本地优先级（简化）

路由器 BGP 配置代码如下。

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 12.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 13.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 1
R1(config-router)#neighbor 13.1.1.2 remote-as 1
R1(config-router)#exit
R1(config)#ip route 24.1.1.0 255.255.255.0 12.1.1.2
R1(config)#ip route 34.1.1.0 255.255.255.0 13.1.1.2
R1(config)#end
R1#
```

路由器 R2 配置如下：

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 24.1.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#access-list 1 permit 1.1.1.0 0.0.0.255
R2(config)#route-map local_pref permit 10
R2(config-route-map)#match ip address 1
R2(config-route-map)#set local-preference 200
R2(config)#route-map local_pref permit 20
R2(config-route-map)#exit
R2(config)#router bgp 1
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 24.1.1.2 remote-as 2
R2(config-router)#neighbor 24.1.1.2 route-map local_pref in
R2(config-router)#end
R2#
```

路由器 R3 配置如下：

```
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 13.1.1.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface fastethernet 0/1
R3(config-if)#ip address 34.1.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#access-list 1 permit 2.2.2.0 0.0.0.255
R3(config)#access-list 1 permit 3.3.3.0 0.0.0.255
```

```
R3(config)#route-map local_pref permit 10
R3(config-route-map)#match ip address 1
R3(config-route-map)#set local-preference 200
R3(config)#route-map local_pref permit 20
R3(config-route-map)#exit
R3(config)#router bgp 1
R3(config-router)#neighbor 13.1.1.1 remote-as 1
R3(config-router)#neighbor 34.1.1.2 remote-as 2
R3(config-router)#neighbor 34.1.1.2 route-map local_pref in
R3(config-router)#end
R3#
```

路由器 R4 配置如下：

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 24.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface fastethernet 0/1
R4(config-if)#ip address 34.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 0
R4(config-if)#ip address 1.1.1.1 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 1
R4(config-if)#ip address 2.2.2.1 255.255.255.0
R4(config-if)#exit
R4(config)#interface Loopback 2
R4(config-if)#ip address 3.3.3.1 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 2
R4(config-router)#neighbor 24.1.1.1 remote-as 1
R4(config-router)#neighbor 34.1.1.1 remote-as 1
R4(config-router)#network 1.1.1.0 mask 255.255.255.0
R4(config-router)#network 2.2.2.0 mask 255.255.255.0
R4(config-router)#network 3.3.3.0 mask 255.255.255.0
R4(config-router)#end
R4#
```

第 2 步：测试结果。

路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 3, local router ID is 3.3.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop                Metric     LocPrf     Weight Path
```

```
*> 1.1.1.0/24        0.0.0.0          0             32768    i
*> 2.2.2.0/24        0.0.0.0          0             32768    i
*> 3.3.3.0/24        0.0.0.0          0             32768    i
Total number of prefixes 3
```

路由器 R4 的 IP 路由表如下：

```
R4#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
C       1.1.1.0/24 is directly connected, Loopback 0
C       1.1.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
C       3.3.3.0/24 is directly connected, Loopback 2
C       3.3.3.1/32 is local host.
C       24.1.1.0/24 is directly connected, FastEthernet 0/0
C       24.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.2/32 is local host.
```

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 3, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf  Path
*> 1.1.1.0/24       34.1.1.2                 0                 2 i
*> 2.2.2.0/24       34.1.1.2                 0        200      2 i
*> 3.3.3.0/24       34.1.1.2                 0        200      2 i
Total number of prefixes 3
```

上述显示结果表明，1.1.1.0/24 网络的 BGP 路由信息本地优先级为默认值；2.2.2.0/24、3.3.3.0/24 网络的 BGP 路由信息本地优先级为 200。R3 的 IP 路由表如下：

```
R3#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
```

```
Gateway of last resort is no set
B       1.1.1.0/24 [20/0] via 34.1.1.2, 00:00:32
B       2.2.2.0/24 [20/0] via 34.1.1.2, 00:00:32
B       3.3.3.0/24 [20/0] via 34.1.1.2, 00:00:32
C       13.1.1.0/24 is directly connected, FastEthernet 0/0
C       13.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 24.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 1.1.1.0/24       24.1.1.2                 0        200          0 2 i
*> 2.2.2.0/24       24.1.1.2                 0                     0 2 i
*> 3.3.3.0/24       24.1.1.2                 0                     0 2 i
Total number of prefixes 3
```

上述显示结果表明，1.1.1.0/24 网络的 BGP 路由信息本地优先级为 200；2.2.2.0/24、3.3.3.0/24 网络的 BGP 路由信息本地优先级为默认值。R3 的 IP 路由表如下：

```
R2#show ip route
Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [20/0] via 24.1.1.2, 00:00:18
B       2.2.2.0/24 [20/0] via 24.1.1.2, 00:00:18
B       3.3.3.0/24 [20/0] via 24.1.1.2, 00:00:18
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
C       24.1.1.0/24 is directly connected, FastEthernet 0/1
C       24.1.1.1/32 is local host.
```

路由器 R4 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 5, local router ID is 13.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
```

```
*>i1.1.1.0/24        24.1.1.2          0      200      0    2 i
* i                  34.1.1.2          0      100      0    2 i
* i2.2.2.0/24        24.1.1.2          0      100      0    2 i
*>i                  34.1.1.2          0      200      0    2 i
* i3.3.3.0/24        24.1.1.2          0      100      0    2 i
*>i                  34.1.1.2          0      200      0    2 i
Total number of prefixes 3
```

上述显示结果表明，路由器 R1 分别从 R2、R3 学到 1.1.1.0/24、2.2.2.0/24、3.3.3.0/24 网络的 BGP 路由信息。但是从 R2 学到的 1.1.1.0/24 网络路由本地优先级为 200，从 R3 学到的 1.1.1.0/24 网络路由本地优先级为默认值 100；同理，从 R2 学到的 2.2.2.0/24、3.3.3.0/24 网络的本地优先级为默认值 100，从 R3 学到的为 200。R1 将本地优先级为 200 的路由加入到 IP 路由表中，R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B     1.1.1.0/24 [200/0] via 24.1.1.2, 00:01:50
B     2.2.2.0/24 [200/0] via 34.1.1.2, 00:01:59
B     3.3.3.0/24 [200/0] via 34.1.1.2, 00:01:59
C     12.1.1.0/24 is directly connected, FastEthernet 0/0
C     12.1.1.1/32 is local host.
C     13.1.1.0/24 is directly connected, FastEthernet 0/1
C     13.1.1.1/32 is local host.
S     24.1.1.0/24 [1/0] via 12.1.1.2
S     34.1.1.0/24 [1/0] via 13.1.1.2
```

从上述 IP 路由表可以看出，去往 1.1.1.0/24 网络的数据流经 R2 到达 R4；去往 2.2.2.0/24、3.3.3.0/24 网络的数据流经 R3 到达 R4，实现了负载均衡，提高了 R2 及 R3 的利用率。

（5）AS 间的度量值属性。

AS 间的度量值属性是可选的非传递属性。有时也称为多出口鉴别器属性（Multi-Exit Discriminators，MED）或直接称为度量值属性。它是影响 BGP 的路径选择的属性。

MED 属性与本地优先级属性不同，MED 属性是可以在 AS 间传递的，也就是说，MED 属性可以被发送到 EBGP 对等体。当一个 EBGP 对等体收到 MED 属性后，会将其传递给 IBGP 对等体，但是该路由的 MED 属性值不会再被通告给另一个自治系统。

在如图 12.28 所示的网络中，路由器 R1 将设置 MED 属性值的路由通告给 R2，此时 MED 属性值也被传递到 R2。R2 将该路由通过 IBGP 通告给 R3 时仍然携带 MED 属性值。但是当 R3 再将此路由通告给 EBGP 邻居 R4 时，MED 属性值将不被传递。

通过前面的学习我们已经知道了本地优先级属性是如何影响 AS 内数据包选择离开自治系统内路径的。反之，MED 属性是影响 AS 外的数据包如何进入本自治系统的。由于 MED 属性值是能够在 AS 之间传递，通过修改 MED 属性值，然后再将修改后的 MED 属性值传递到其他 AS，从而影响其他 AS

进入本自治系统的入口选择。

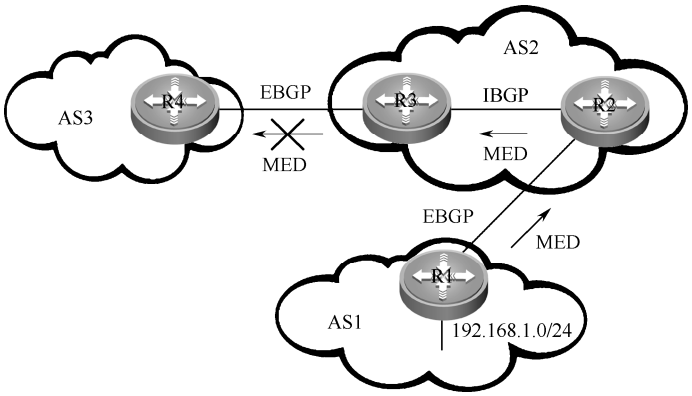


图 12.28　MED 属性传递

第一，定义没有设置 MED 属性值的路由为最差路由。

默认情况下，如果接收到没有设置 MED 属性的路径，该路径的 MED 值被认为是 0。根据 MED 值越小，路径优先级越高的原则，这种路由的优先级最高。如果希望没有设置 MED 属性的路径的优先级为最低，则可以使用 bgp bestpath med missing-as-worst 命令将丢弃 MED 属性值的路由变成最差的路由。该命令具体格式如下：

ruijie(config-router)#bgp bestpath med missing-as-worst

设备名称为ruijie
处于路由配置模式

使用该命令的 no 选项将关闭该功能，如 ruijie(config-router)# no bgp bestpath med missing-as-worst。

例如：将没有设置 MED 属性的 BGP 路径优先级定义为最差的路由。

```
ruijie(config)#router bgp 65000
ruijie(config-router)#bgp bestpath med missing-as-worst
```

第二，使用路由策略配置 MED 属性值。

与配置本地优先级一样，默认 MED 值不能实现差异化路由 MED 值设置。如果要实现差异化配置，必须使用前面介绍过的 route-map 命令。只是在配置 set 语句时使用 set metric metric 命令对 MED 属性值进行设置。该命令具体格式如下：

ruijie(config-route-map)#set metric *metric*

设备名称为ruijie
处于路由图配置模式　　*设置的MED值*

其中，*metric* 为设置的 MED 值，数值越小者拥有更高的优先级。如果对特定的路由进行设置，还需要进行 match 命令配置，以匹配一条或多条路由。该命令的 no 形式删除已有配置，如 ruijie(config-router)# no set metric。

使用 route-map 命令对路由的 MED 值进行定义后，还需要在 BGP 中引用已定义的 route-map 语句。该命令格式如下：

ruijie(config-router)#neighbor *peer-address* **route-map** *map-tag* **out**

设备名称为ruijie 处于路由配置模式 *对等体IP地址* *路由图名称*

命令格式中的参数与之前的全部相同，只是在设置 MED 值时，通常使用 out 对通告给 EBGP 对等体的路由进行设置，引入 MED 属性用来影响 EBGP 对等体的数据流。使用该命令的 no 选项关闭该功能，如 ruijie(config-router)#no reighbor 192.168.12.1 route-map medmap out。

第三，“AS 间的度量值属性”应用实例。

在如图 12.29 所示网络中，希望将去往 1.1.1.0/24 的流量从 R2 进入，去往 2.2.2.0/24、3.3.3.0/24 的流量从 R3 进入。为此，分别在 R2 和 R3 上配置策略路由，对 metric 值进行修改。

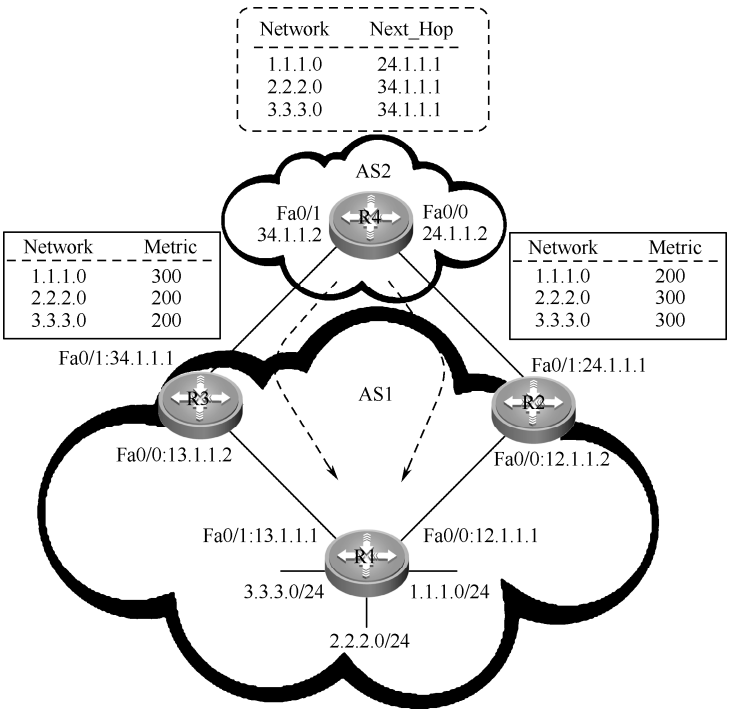


图 12.29　基于策略路由设置 MED 属性值

路由器 BGP 配置代码如下。

第 1 步：BGP 基本配置。

路由器 R1 配置如下：

```
R1(config)#interface fastethernet 0/0
R1(config-if)#ip address 12.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface fastethernet 0/1
R1(config-if)#ip address 13.1.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface Loopback 0
R1(config-if)#ip address 1.1.1.1 255.255.255.0
R1(config-if)#exit
```

```
R1(config)#interface Loopback 1
R1(config-if)#ip address 2.2.2.1 255.255.255.0
R1(config-if)#exit
R1(config)#interface Loopback 2
R1(config-if)#ip address 3.3.3.1 255.255.255.0
R1(config-if)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 1
R1(config-router)#neighbor 13.1.1.2 remote-as 1
R1(config-router)#network 1.1.1.0 mask 255.255.255.0
R1(config-router)#network 2.2.2.0 mask 255.255.255.0
R1(config-router)#network 3.3.3.0 mask 255.255.255.0
R1(config)#end
R1#
```

路由器 R2 配置如下：

```
R2(config)#interface fastethernet 0/0
R2(config-if)#ip address 12.1.1.2 255.255.255.0
R2(config-if)#exit
R2(config)#interface fastethernet 0/1
R2(config-if)#ip address 24.1.1.1 255.255.255.0
R2(config-if)#exit
R2(config)#access-list 1 permit 1.1.1.0 0.0.0.255
R2(config)#route-map med permit 10
R2(config-route-map)#match ip address 1
R2(config-route-map)#set set metric 200
R2(config)#route-map med permit 20
R2(config-route-map)#set set metric 300
R2(config-route-map)#exit
R2(config)#router bgp 1
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 24.1.1.2 remote-as 2
R2(config-router)#neighbor 24.1.1.2 route-map med out
R2(config-router)#end
R2#
```

路由器 R3 配置如下：

```
R3(config)#interface fastethernet 0/0
R3(config-if)#ip address 13.1.1.2 255.255.255.0
R3(config-if)#exit
R3(config)#interface fastethernet 0/1
R3(config-if)#ip address 34.1.1.1 255.255.255.0
R3(config-if)#exit
R3(config)#access-list 1 permit 2.2.2.0 0.0.0.255
R3(config)#access-list 1 permit 3.3.3.0 0.0.0.255
R3(config)#route-map med permit 10
```

```
R3(config-route-map)#match ip address 1
R3(config-route-map)#set metric 200
R3(config)#route-map med permit 20
R3(config-route-map)#set metric 300
R3(config-route-map)#exit
R3(config)#router bgp 1
R3(config-router)#neighbor 13.1.1.1 remote-as 1
R3(config-router)#neighbor 34.1.1.2 remote-as 2
R3(config-router)#neighbor 34.1.1.2 route-map med out
R3(config-router)#end
R3#
```

路由器 R4 配置如下：

```
R4(config)#interface fastethernet 0/0
R4(config-if)#ip address 24.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#interface fastethernet 0/1
R4(config-if)#ip address 34.1.1.2 255.255.255.0
R4(config-if)#exit
R4(config)#router bgp 2
R4(config-router)#neighbor 24.1.1.1 remote-as 1
R4(config-router)#neighbor 34.1.1.1 remote-as 1
R4(config-router)#end
R4#
```

第 2 步：测试结果。

路由器 R1 的 BGP 路由表如下：

```
R1#show ip bgp
BGP table version is 1, local router ID is 13.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*> 1.1.1.0/24       0.0.0.0                  0                    32768 i
*> 2.2.2.0/24       0.0.0.0                  0                    32768 i
*> 3.3.3.0/24       0.0.0.0                  0                    32768 i
Total number of prefixes 3
```

从上述显示结果看出，R1 本地产生 3 条 BGP 路由信息。

路由器 R1 的 IP 路由表如下：

```
R1#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
```

```
Gateway of last resort is no set
C       1.1.1.0/24 is directly connected, Loopback 0
C       1.1.1.1/32 is local host.
C       2.2.2.0/24 is directly connected, Loopback 1
C       2.2.2.1/32 is local host.
C       3.3.3.0/24 is directly connected, Loopback 2
C       3.3.3.1/32 is local host.
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.1/32 is local host.
C       13.1.1.0/24 is directly connected, FastEthernet 0/1
C       13.1.1.1/32 is local host.
```

路由器 R2 的 BGP 路由表如下：

```
R2#show ip bgp
BGP table version is 3, local router ID is 24.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf     Weight Path
*>i1.1.1.0/24       12.1.1.1                 0        100          0 i
*>i2.2.2.0/24       12.1.1.1                 0        100          0 i
*>i3.3.3.0/24       12.1.1.1                 0        100          0 i
Total number of prefixes 3
```

从上述显示结果看出，R2 从 R1 学习到 3 条 BGP 路由信息。

路由器 R2 的 IP 路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [200/0] via 12.1.1.1, 00:03:40
B       2.2.2.0/24 [200/0] via 12.1.1.1, 00:03:40
B       3.3.3.0/24 [200/0] via 12.1.1.1, 00:03:40
C       12.1.1.0/24 is directly connected, FastEthernet 0/0
C       12.1.1.2/32 is local host.
C       24.1.1.0/24 is directly connected, FastEthernet 0/1
C       24.1.1.1/32 is local host.
```

从上述显示结果可以看出，有 3 条 BGP 路由信息加入到 IP 路由表中。

路由器 R3 的 BGP 路由表如下：

```
R3#show ip bgp
BGP table version is 3, local router ID is 34.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
```

```
                S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric     LocPrf  Path
*>i1.1.1.0/24       13.1.1.1                 0        100  i
*>i2.2.2.0/24       13.1.1.1                 0        100  i
*>i3.3.3.0/24       13.1.1.1                 0        100  i
Total number of prefixes 3
```

从上述显示结果看出，路由器 R2 从 R1 学习到 3 条 BGP 路由信息。

路由器 R3 的 IP 路由表如下：

```
R3#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [200/0] via 13.1.1.1, 00:02:34
B       2.2.2.0/24 [200/0] via 13.1.1.1, 00:02:34
B       3.3.3.0/24 [200/0] via 13.1.1.1, 00:02:34
C       13.1.1.0/24 is directly connected, FastEthernet 0/0
C       13.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.1/32 is local host.
```

从上述显示结果可以看出，有 3 条 BGP 路由信息加入到 IP 路由表中。

路由器 R4 的 BGP 路由表如下：

```
R4#show ip bgp
BGP table version is 3, local router ID is 34.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
                S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
    Network          Next Hop            Metric     LocPrf     Weight Path
*   1.1.1.0/24       34.1.1.1                 300                    0 1 i
*>                   24.1.1.1                 200                    0 1 i
*> 2.2.2.0/24        34.1.1.1                 200                    0 1 i
*                    24.1.1.1                 300                    0 1 i
*> 3.3.3.0/24        34.1.1.1                 200                    0 1 i
*                    24.1.1.1                 300                    0 1 i
Total number of prefixes 3
```

从上述显示结果看出，从 R2 学习到 BGP 路由中，去往 1.1.1.0/24 网络路由的 metric 值为 200，其余路由的 metric 值为 300；从 R3 学习到 BGP 路由中，去往 2.2.2.0/24、3.3.3.0/24 网络路由的 metric 值为 200，其余路由的 metric 值为 300。

路由器 R4 的 IP 路由表如下：

```
R4#show ip route
Codes:  C - connected, S - static, R - RIP, B - BGP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default
Gateway of last resort is no set
B       1.1.1.0/24 [20/200] via 24.1.1.1, 00:00:26
B       2.2.2.0/24 [20/200] via 34.1.1.1, 00:00:26
B       3.3.3.0/24 [20/200] via 34.1.1.1, 00:00:26
C       24.1.1.0/24 is directly connected, FastEthernet 0/0
C       24.1.1.2/32 is local host.
C       34.1.1.0/24 is directly connected, FastEthernet 0/1
C       34.1.1.2/32 is local host.
```

从上述显示结果可以看出，将 metric 为 200 的最佳路由加入到 IP 路由表中。

3. BGP 路由决策规则

上述内容介绍了 BGP 路由协议的各种属性，其中大部分属性对路径决策起到决定作用。在上述介绍的多个属性都存在的情况下，如何比较这些属性值以决定最佳路径呢？具体比较过程如下。

（1）比较本地优先级，选举出本地优先级最高（数值大）的路由。

（2）若本地优先级相同，则优先选择本地路由器通告的路由。

（3）若没有本地路由器通告的路由，则比较 AS 路径列表，选择 AS 路径最短的路由。

（4）若 AS 路径列表的长度相同，则比较起源属性。起源属性的优先级从高到低顺序为 IGP>EGP>Incomplete。

（5）若起源属性相同，则比较 MED 值，选择最小 MED 值的路由。

（6）若 MED 值相同，则比较邻居关系，邻居关系优先级为 EBGP > IBGP，优选 EBGP。

（7）若邻居关系相同，则比较达到下一跳的 IGP 度量值，优选到达下一跳的 IGP 度量值最小的路由。

（8）若路径都是 EBGP 路由，则优选最旧的，也就是获得时间最长的路由。

（9）优选邻居路由器 ID 最小的路由。

（10）若邻居路由器 ID 相同，则优选邻居 IP 最小的路由。

经过以上过程，BGP 从多条到达相同目的地的路由中选择出一条最佳路由，然后将该路由添加到 IP 路由表中，并将最佳路径通告给其他 BGP 对等体。

练习题

1. 选择题

（1）BGP 协议使用的传输协议、源接口号、目标接口号是（　　）。

A．TCP　179　　B．TCP　20　　C．UDP　67　　D．UDP　68

（2）下面关于 BGP 路由协议的描述错误的是（　　）。

A．每个运行 BGP 协议的路由器都会建立一个 BGP 数据库，存在本地 AS 和从其他 AS 接收

的 BGP 路由信息

B．BGP 是运行在 AS 之间的外部网关路由协议

C．BGP 路由协议依据一系列路径属性判定路径的优劣

D．BGP 与 RIP 一样，也是一个距离向量路由协议

2. 简答题

（1）举例说明 BGP 协议中 AS 列表属性的作用。

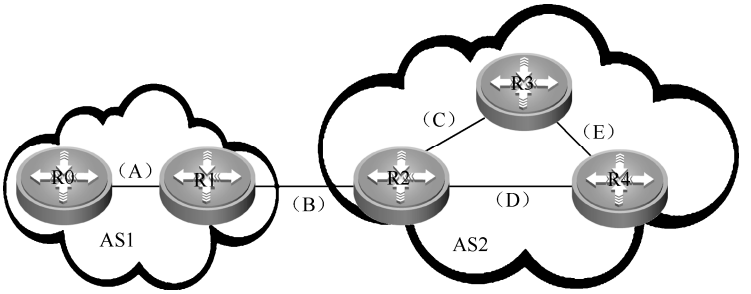
（2）举例说明 BGP 是如何防止路由环路的。

（3）举例说明什么是 IBGP 及 EBGP。

（4）在 BGP 路由表中的起源属性“i”、“e”、“？”分别表示什么？

3. 填空题

（1）在下面网络拓扑图中，AS1 与 AS2 中所有路由器启用 BGP 路由协议，请指出路由器之间 BGP 对等体之间连接关系。



（A）在标记“A”处 BGP 对等体关系是_____。

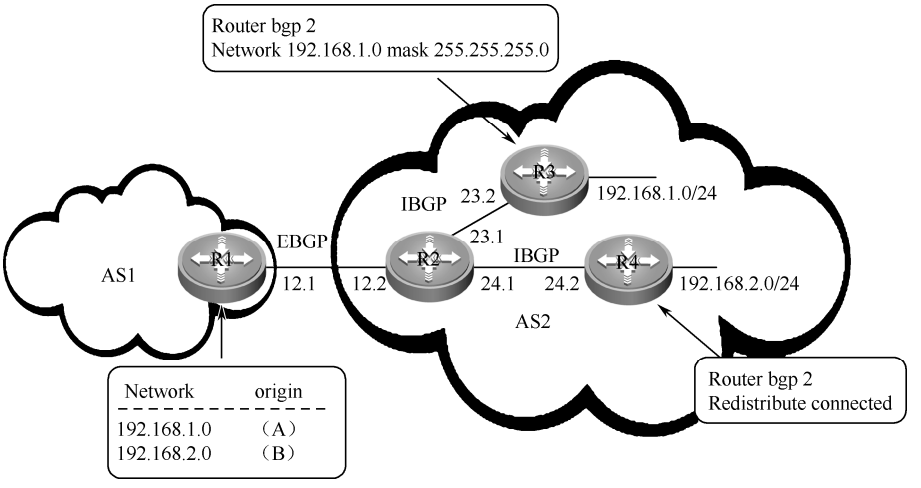
（B）在标记“B”处 BGP 对等体关系是_____。

（C）在标记“C”处 BGP 对等体关系是_____。

（D）在标记“D”处 BGP 对等体关系是_____。

（E）在标记“E”处 BGP 对等体关系是_____。

（2）根据下面网络拓扑图，回答相关提问。



（A）在路由器 R1 中 192.168.1.0 路由的起源属性是_____。

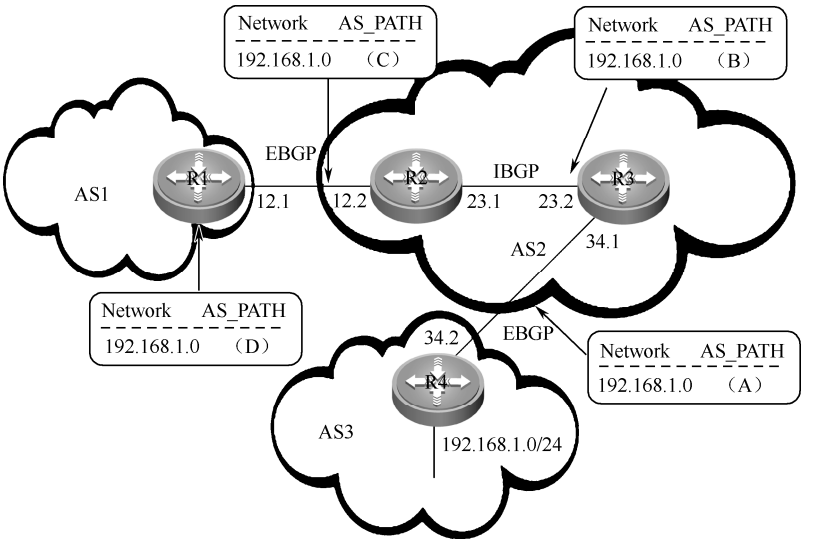
（B）在路由器 R1 中 192.168.2.0 路由的起源属性是 _____。

（C）在路由器 R1 上配置路由器 R2 为 BGP 邻居的命令为_____。

（D）在路由器 R1 上获得 192.168.1.0 路由的下一跳地址是_____。

（E）在路由器 R1 上获得 192.168.2.0 路由的下一跳地址是_____。

（3）在下面网络拓扑图中，根据路由行走路径分别填写 AS 列表及回答相关提问。



（A）标记“A”处的 AS 列表为_____。

（B）标记“B”处的 AS 列表为_____。

（C）标记“C”处的 AS 列表为_____。

（D）标记“D”处的 AS 列表为_____。

（E）在路由器 R3 上获得 192.168.1.0 路由的下一跳地址是_____。

项目 13　GRE 隧道协议

在网络工程中，经常遇到分布在不同地理位置的总公司与分公司通过 ISP 网络连接起来，实现公司内部资源共享。可以在 ISP 公共网络中建立一条隧道，实现总公司与分公司路由信息交换，达到相互通信的目的。本章介绍 GRE 工作原理及相关配置命令使用。

知识点、技能点

1. 了解 GRE 工作原理;
2. 掌握 GRE 配置技能。

13.1　问题提出

通过公共网络将总公司与分公司连接起来，如何实现全网络路由互通，可以通过建立 GRE 隧道，实现网络路由信息交换。

13.2　相关知识

1. GRE 工作原理

GRE（Generic Routing Encapsulation，通用路由协议封装）是由 Cisco 和 Net-Smiths 等公司于 1994 年提交给 IETF 的，标号为 RFC1701 和 RFC1702，用于传输三层网络协议的隧道协议。目前有多数厂商的网络设备均支持 GRE 隧道协议。

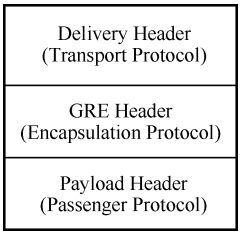


图 13.1　GRE 传输报文格式

路由器收到一个需要封装和路由的原始数据报文（Payload），这个报文首先被 GRE 封装成 GRE 报文，然后又被封装成 IP 报文，由 IP 协议负责此报文的转发。

原始报文的协议被称为乘客协议，GRE 被称为封装协议，而负责转发的 IP 协议被称为传输协议，如图 13.1 所示。

GRE 规定了如何用一种网络协议去封装另一种网络协议的方法。GRE 的隧道由两端的源 IP 地址和目的 IP 地址来定义，允许用户使用 IP 包封装 IP、IPX、AppleTalk 包，并支持全部的路由协议（如 RIP2、OSPF 等）。通过 GRE，用户可以利用公共 IP 网络连接 IPX 网络、AppleTalk 网络，还可以使用保留地址进行网络互连，或者对公网隐藏企业网的 IP 地址，如图 13.2 所示。

GRE 只提供了数据包的封装，它并没有加密功能来防止网络侦听和攻击。所以在实际环境中它常和 IPSec 一起使用，由 IPSec 提供用户数据的加密，从而给用户提供更好的安全性。

GRE 具有的优点如下。

（1）多协议的本地网络可以通过单一协议的骨干网实现传输。

（2）将一些不能连续的子网连接起来，用于组建 VPN。

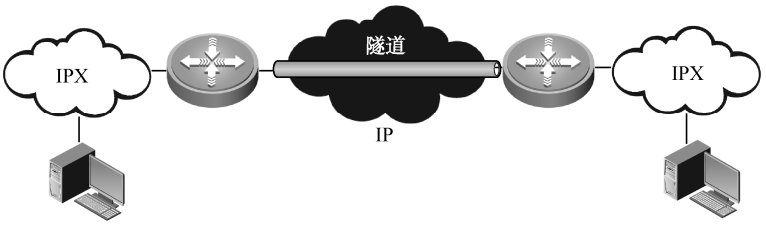


图 13.2　GRE 建立隧道

（3）扩大网络的工作范围。如 RIP 最多 16 跳，GRE 连接隧道计 1 跳。

2. GRE 配置命令

GRE 协议主要应用在隧道（Tunnel）接口中。路由器通过 Tunnel 接口来实现隧道功能，允许利用传输协议（如 IP）来传送任意协议的网络数据包。同其他逻辑接口一样，Tunnel 接口也是系统虚拟的接口。Tunnel 接口并不特别指定传输协议或者负载协议，它提供的是一个用来实现标准的点对点的传输模式。由于 Tunnel 实现的是点对点的传输链路，所以对于每一个单独的链路都必须设置一个 Tunnel 接口。

Tunnel 功能实现包括下面三个主要组成部分。

第一，乘客协议，就是通过 Tunnel 传输的负载（网络数据）的封装协议。目前锐捷产品软件在 Tunnel 接口上只支持 IP 协议作为乘客协议（负载协议）。

第二，封装协议，就是用来二次封装并辨识待传输负载的协议。锐捷产品在 Tunnel 接口上支持的封装模式有 GRE、IPIP。

第三，传输协议，就是实际传输经过载体协议二次封装后的网络协议。锐捷产品使用最广泛应用的 IP 协议作为传输协议。

（1）创建 Tunnel 接口，并进入接口配置模式，使用的命令格式如下：

指定接口类型为Tunnel

ruijie(config)#interface tunnel *tunnel-number*

设备名称为ruijie
处于全局配置模式　　*接口编号*

其中，*tunnel-number* 为 Tunnel 接口编号，范围为 1～255。同其他逻辑接口一样，第一次进入指定的 Tunnel 接口的时候就创建了一个 Tunnel 接口。

例如：创建隧道接口 Tunnel 1，并设置该接口 IP 地址为 192.168.1.1。

```
ruijie(config)#interface tunnel 1
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
```

（2）设置 Tunnel 接口源地址。一个 Tunnel 接口需要明确配置隧道的源地址和目的地址，为了保证隧道接口的稳定性，一般将 loopback 地址作为隧道的源地址和目的地址。在 Tunnel 接口正常工作之前，需要确认源地址和目的地址的连通性。

要配置 Tunnel 接口的源地址，在 Tunnel 接口配置模式中执行以下命令：

用IP地址方式指定Tunnel 源地址

ruijie(config-if)#tunnel source { *ip-address* | *interface-type interface-number* }

设备名称为ruijie
处于接口配置模式　　*用接口方式指定Tunnel源地址*

其中，*ip-address* 为用来指定 Tunnel 源地址的 IP 地址，在设备上已经设置好的其他接口的 IP 地址；*interface-type* 为通用的接口类型，如以太网接口 Ethernet、快速以太网接口 fastEthernet、loopback 接口、null 接口以及其他 Tunnel 接口等；*interface-number* 为接口编号。

此处设置 Tunnel 接口源地址是 Tunnel 接口用来进行实际通信的源地址，它也是 Tunnel 位于本地的端点。

默认情况下，没有配置源地址。使用本命令的 no 形式删除 Tunnel 接口的源地址，如 ruijie(config-if)# no tunnel source。

例如：在 Tunnel 0 接口上设置命令 tunnel source 指定串口 2/0 为源地址接口。

```
ruijie(config)#interface tunnel 0
ruijie(config-if)#tunnel source serial 2/0
```

（3）设置 Tunnel 接口目的地址。要设置 Tunnel 接口的目的地址，在 Tunnel 接口配置模式中执行以下命令：

指定Tunnel目的地址

ruijie(config-if)#tunnel destination *ip-address*

设备名称为ruijie
处于接口配置模式

其中，*ip-address* 为被指定的隧道目标的 IP 地址。默认情况下，没有配置目的 IP 地址。使用本命令的 no 形式删除 Tunnel 接口的目标地址，如 ruijie(config-if)# no tunnel destination。

此处设置 Tunnel 接口目的地址是 Tunnel 接口用来进行实际通信的目的地址，它也是 Tunnel 位于远程对端的端点。

例如：在 Tunnel 0 接口上配置目标的 IP 地址 61.154.101.3。

```
ruijie(config)#interface tunnel 0
ruijie(config-if)#tunnel destination 61.154.101.3
```

（4）设置 Tunnel 封装格式（可选）。Tunnel 的封装格式就是 Tunnel 的载体协议。Tunnel 接口的默认封装格式是 GRE。当然，用户也可以根据实际使用情况来决定 Tunnel 接口的封装格式，在默认的情况下，可以不对封装格式做任何的定义就可以实现 IP 隧道的 GRE 封装。设置 Tunnel 封装格式的命令如下：

指定Tunnel封装形式

ruijie(config-if)#tunnel mode *{ gre | ipip }*

设备名称为ruijie
处于接口配置模式

其中，*gre* 为可以架构在 IP 层的通用路由封装 GRE 协议；*ipip* 为 IP over IP 的封装模式。

默认情况下，使用 GRE 封装。使用本命令的 no 形式恢复默认方式。

例如：在 Tunnel 0 接口上封装 IPIP 协议。

```
ruijie(config)#interface tunnel 0
ruijie(config-if)#tunnel mode ipip
```

（5）设置认证密码（可选）。隧道的建立如果没有一些密钥保护，可能会受到非法的入侵或者报文攻击。使用 tunnel key 命令只有在 GRE 封装才可以生效，该命令格式如下：

指定Tunnel认证密码

ruijie(config-if)#tunnel key *value*

设备名称为ruijie
处于接口配置模式

其中，*value* 为 Tunnel 的密钥值，取值范围在 0～4294967295。默认情况下，没有配置认证密码。使用本命令的 no 形式删除该密钥，如 ruijie(config-if)# no tunnel key。

例如：在 Tunnel 0 接口中设置密钥 1234。

```
ruijie(config)#interface tunnel 0
ruijie(config-if)#tunnel key 1234
```

（6）查看 Tunnel 接口配置情况。为了查看 Tunnel 接口状态，使用如下命令：

Tunnel接口编号

ruijie#show interfaces tunnel *tunnel-number*

设备名称为ruijie
处于特权模式

其中，*tunnel-number* 为 tunnel 接口编号。

例如：显示 Tunnel 1 接口的状态。

```
ruijie#show interfaces tunnel 1
Tunnel 1 is UP , line protocol is UP
Hardware is Tunnel
Interface address is: 1.1.1.1/24
MTU 1500 bytes, BW 9 Kbit
Encapsulation protocol is Tunnel, loopback not set
Keepalive interval is 0 sec , no set
Carrier delay is 0 sec
RXload is 1 ,Txload is 1
Tunnel source 192.168.200.200 (FastEthernet 0/0), destination
192.168.200.100
Tunnel protocol/transport GRE/IP, key 0xea
Order sequence numbers 0/0 (tx/rx)
Checksumming of packets enabled Queueing strategy: WFQ
5 minutes input rate 0 bits/sec, 0 packets/sec
5 minutes output rate 0 bits/sec, 0 packets/sec
0 packets input, 0 bytes, 0 no buffer
Received 0 broadcasts, 0 runts, 0 giants
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
0 packets output, 0 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
```

从上面的显示结果可以了解 Tunnel 接口的参数设置及工作的情况，比如接口与链路状态、IP 地址设置、MTU 设置、带宽 Bandwidth 设置等。

3. GRE 配置实例

在如图 13.3 所示网络中，路由器 R1 与 R2 之间建立 Tunnel，路由器 R1 背后的子网 202.126.101.0/24 与路由器 R2 背后的子网 67.151.69.0/24 通过 R1 与 R2 之间的 Tunnel 进行通信。这种通信通过 Tunnel 进行，对于 R1 与 R2 之间的外部网络是透明的和不可见的，即是一种虚拟专用网（VPN）的实现。下面将给出路由器 R1 与 R2 的有关 Tunnel 的相关配置。

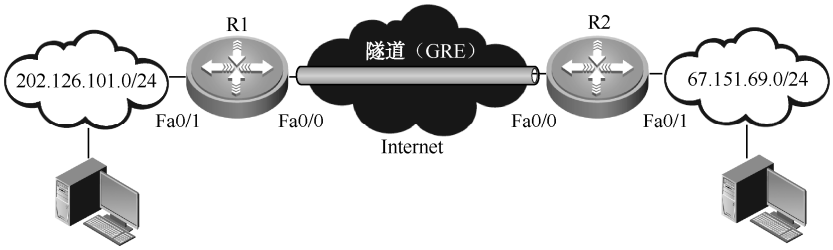


图 13.3　GRE 应用实例（一）

（1）路由器 R1 的相关配置：

```
R1(config)#interface Tunnel0
R1(config-if)#ip address 21.21.21.3 255.255.255.0
R1(config-if)#tunnel source 179.208.12.221
R1(config-if)#tunnel destination 179.208.12.55
R1(config-if)#exit
R1(config)#interface FastEthernet0/0
R1(config-if)#ip address 179.208.12.221 255.255.255.0
R1(config-if)#exit
R1(config)#interface FastEthernet0/1
R1(config-if)#ip address 202.106.101.2 255.255.255.0
R1(config-if)#exit
R1(config)#
```

（2）路由器 R2 的相关配置：

```
R2(config)#interface Tunnel0
R2(config-if)#ip address 21.21.21.5 255.255.255.0
R2(config-if)#tunnel source 179.208.12.55
R2(config-if)#tunnel destination 179.208.12.221
R2(config-if)#exit
R2(config)#interface FastEthernet0/0
R2(config-if)#ip address 179.208.12.55 255.255.255.0
R2(config-if)#exit
R2(config)#interface FastEthernet0/1
R2(config-if)#ip address 67.151.69.202 255.255.255.0
R2(config-if)#exit
R2(config)#
```

从上面的配置可以看出，路由器 R1 与 R2 均使用以太网接口 Fa0/0 建立 Tunnel，使用以太网接口 Fa0/1 与内部网连接并作为内部网的网关。

13.3　扩展知识

1. Loopback 接口

Loopback（回环）接口是完全由软件模拟的设备本地接口，它永远都处于 UP 状态。发往 Loopback 接口的数据包将会在设备本地处理，包括路由信息。Loopback 接口的 IP 地址可以用来作为 OSPF 路由协议的设备标志、实施发向 Telnet 或者作为远程 Telnet 访问的网络接口等。配置一个 Loopback 接口类似于配置一个以太网接口，可以把它看作一个虚拟的以太网接口。

创建 loopback 接口命令如下：

```
ruijie(config)#interface loopback loopback-interface-number
```

其中，*loopback-interface-number* 为接口编号，范围为 0～255。

使用命令 interface loopback loopback-interface-number 创建一个指定接口号的 Loopback 后，就可以像配置以太网口一样配置 Loopback 接口的通信参数（如 IP 地址等）。

由于 Loopback 是虚拟的接口，它只在逻辑意义上存在，可以在需要的时候使用命令 no interface loopback loopback-interface-number 来删除指定的 Loopback 接口。系统提供命令 show interfaces loopback loopback-interface-number 来监控与维护 Loopback 接口。

例如：创建 Loopback 0 接口，并设置该接口 IP 地址为 192.168.1.1。

```
ruijie(config)#interface loopback 0
ruijie(config-if)#ip address 192.168.1.1 255.255.255.0
```

2. NULL 接口

路由器还提供另外一个逻辑接口，即 NULL（空）的虚拟接口。该虚拟接口仅仅相当于一个可用的系统设备。NULL 永远都处于 UP 状态并且永远不会主动发送或者接受网络数据，任何发往 NULL 接口的数据包都会被丢弃。在 NULL 接口上不能配置任何命令（不包括每一个接口都有的命令 help 与 exit）。

NULL 接口更多地用于网络数据流的过滤。如果使用空接口，可以通过将不希望处理的网络数据流路由给 NULL 接口而不必使用访问列表。

例如：将去往 127.0.0.0/8 网络的流量过滤掉。

```
ruijie(config)#ip route 127.0.0.0 255.0.0.0 null 0
```

由于 NULL 接口在功能上是一个“空”的系统设备，作为系统设备是无法使用 no interface null 命令来删除 NULL 接口的。

13.4　演示实例

1. 项目描述

北京的某公司在上海设立了新的分公司，分公司要能够访问总公司的各种网络资源，并且要求分公司和总公司之间共享路由信息。这家公司希望通过 VPN 技术实现两个站点的数据传输。

为了实现北京总公司与上海分公司之间通过 Internet 进行路由信息和数据信息传输的问题。需要采用 GRE VPN 技术，通过隧道技术有效地保证数据在 Internet 网络的传输，并且 GRE 支持对组播和广播数据的封装，可用于封装路由协议报文，如图 13.4 所示。

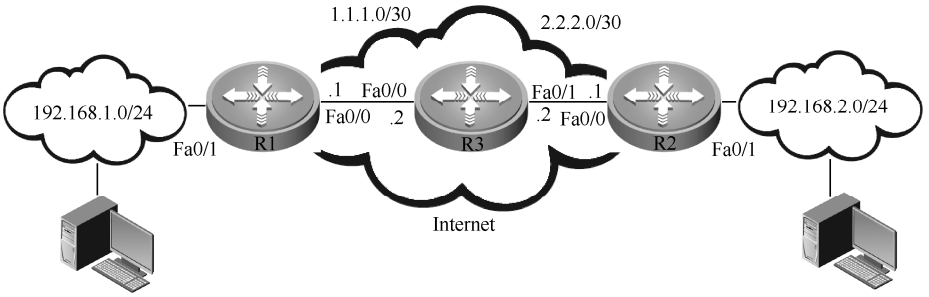


图 13.4 GRE 应用实例（二）

2. 操作步骤

第 1 步：配置 Internet 路由器 R3。

```
R3#configure terminal
R3(config)#interface fastEthernet 0/0
R3(config-if)#ip address 1.1.1.2 255.255.255.252
R3(config-if)#exit
R3(config)#interface fastEthernet 0/1
R3(config-if)#ip address 2.2.2.2 255.255.255.252
R3(config-if)#exit
```

第 2 步：配置路由器 R1 与 R2 的 Internet 连通性。

（1）配置路由器 R1。

```
R1#configure terminal
R1(config)#interface fastEthernet 0/0
R1(config-if)#ip address 1.1.1.1 255.255.255.252
R1(config-if)#exit
R1(config)#interface fastEthernet 0/1
R1(config-if)#ip address 192.168.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#ip route 0.0.0.0 0.0.0.0 1.1.1.2
```

（2）配置路由器 R2。

```
R2#configure terminal
R2(config)#interface fastEthernet f0/0
R2(config-if)#ip address 2.2.2.1 255.255.255.252
R2(config-if)#exit
R2(config)#interface fastEthernet 0/1
R2(config-if)#ip address 192.168.2.1 255.255.255.0
R2(config-if)#exit
R2(config)#ip route 0.0.0.0 0.0.0.0 2.2.2.2
```

第 3 步：配置路由器 R1 的 GRE 隧道。

```
R1(config)#interface tunnel 1
R1(config-if)#ip address 10.1.1.1 255.255.255.0
R1(config-if)#tunnel source fastEthernet 0/0
R1(config-if)#tunnel destination 2.2.2.1
```

```
R1(config-if)#tunnel key 1234567
R1(config-if)#exit
```

第 4 步：在路由器 R1 上启用 RIPv2 路由协议。

```
R1(config)#router rip
R1(config-router)#version 2
R1(config-router)#no auto-summary
R1(config-router)#network 10.0.0.0
R1(config-router)#network 192.168.1.0
R1(config-router)#exit
```

第 5 步：配置路由器 R2 的 GRE 隧道。

```
R2(config)#interface tunnel1
R2(config-if)#ip address 10.1.1.2 255.255.255.0
R2(config-if)#tunnel source fastEthernet 0/0
R2(config-if)#tunnel destination 1.1.1.1
R2(config-if)#tunnel key 1234567
R2(config-if)#exit
```

第 6 步：在路由器 R2 上启用 RIPv2 路由协议。

```
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#no auto-summary
R2(config-router)#network 10.0.0.0
R2(config-router)#network 192.168.2.0
R2(config-router)#exit
```

第 7 步：配置主机 PC1 和 PC2。

PC1 的 IP 地址为 192.168.1.2，网关为 192.168.1.1。

PC2 的 IP 地址为 192.168.2.2，网关为 192.168.2.1。

第 8 步：验证测试路由器的配置。

在路由器 R1 与 R2 上验证 GRE 隧道状态及路由表信息，分别通过 Tunnel 接口学习到对端局域网的路由。

显示路由器 R1 上的 Tunnel 1 接口状态如下：

```
R1#show interface tunnel 1
Tunnel 1 is UP, line protocol is UP
Hardware is Tunnel
Interface address is: 10.1.1.1/24
   MTU 1472 bytes, BW 9 Kbit
   Encapsulation protocol is Tunnel, loopback not set
   Keepalive interval is 0 sec , no set
   Carrier delay is 0 sec
   RXload is 1 ,Txload is 1
   Tunnel source 1.1.1.1 (FastEthernet 1/0), destination 2.2.2.1
   Tunnel protocol/transport GRE/IP, key 0x12d687, sequencing disabled
   Checksumming of packets disabled    Queueing strategy: WFQ
   5 minutes input rate 0 bits/sec, 0 packets/sec
```

```
    5 minutes output rate 12 bits/sec, 0 packets/sec
       0 packets input, 0 bytes, 0 no buffer
       Received 0 broadcasts, 0 runts, 0 giants
       0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
       19 packets output, 988 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
```

显示路由器 R1 的路由表如下：

```
R1#show ip route
Codes:   C - connected, S - static,    R - RIP
            O - OSPF, IA - OSPF inter area
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            * - candidate default
Gateway of last resort is 1.1.1.2 to network 0.0.0.0
S*      0.0.0.0/0 [1/0] via 1.1.1.2
C       1.1.1.0/30 is directly connected, FastEthernet 1/0
C       1.1.1.1/32 is local host.
C       10.1.1.0/24 is directly connected, Tunnel 1
C       10.1.1.1/32 is local host.
C       192.168.1.0/24 is directly connected, FastEthernet 1/1
C       192.168.1.1/32 is local host.
R       192.168.2.0/24 [120/1] via 10.1.1.2, 00:00:29, Tunnel 1
```

显示路由器 R2 上的 Tunnel 1 接口状态如下：

```
R2#show interface tunnel 1
Tunnel 1 is UP, line protocol is UP          ！隧道状态为 UP
Hardware is Tunnel
Interface address is: 10.1.1.2/24
   MTU 1472 bytes, BW 9 Kbit
   Encapsulation protocol is Tunnel, loopback not set
   Keepalive interval is 0 sec , no set
   Carrier delay is 0 sec
   RXload is 1 ,Txload is 1
   Tunnel source 2.2.2.1 (FastEthernet 1/1), destination 1.1.1.1
   Tunnel protocol/transport GRE/IP, key 0x12d687, sequencing disabled
   Checksumming of packets disabled    Queueing strategy: WFQ
   5 minutes input rate 31 bits/sec, 0 packets/sec
   5 minutes output rate 36 bits/sec, 0 packets/sec
      55 packets input, 3700 bytes, 0 no buffer
      Received 0 broadcasts, 0 runts, 0 giants
      0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
      58 packets output, 3980 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
```

显示路由器 R2 的路由表如下：

```
R2#show ip route
Codes:  C - connected, S - static,   R - RIP
        O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        * - candidate default
Gateway of last resort is 2.2.2.2 to network 0.0.0.0
S*    0.0.0.0/0 [1/0] via 2.2.2.2
C     2.2.2.0/30 is directly connected, FastEthernet 1/1
C     2.2.2.1/32 is local host.
C     10.1.1.0/24 is directly connected, Tunnel 1
C     10.1.1.2/32 is local host.
R     192.168.1.0/24 [120/1] via 10.1.1.1, 00:00:20, Tunnel 1
C     192.168.2.0/24 is directly connected, FastEthernet 1/0
C     192.168.2.1/32 is local host.
```

第 9 步：验证测试主机的连通性。

在主机 PC1 上 ping PC2，可以 ping 通。

注意事项：

（1）GRE 隧道两端的密钥要匹配。

（2）隧道两端的源和目的相互对应，即路由器 R1 的源地址为路由器 R2 的目的地址，R2 的源地址为 R1 的目的地址。

（3）需要在 Tunnel 接口启用路由，而非连接 Internet 的接口。

13.5　训练任务

模拟总公司与分公司通过 Internet 连接，为了实现公司内路由信息更新，总公司与分公司建立 GRE 隧道。

要求：

1．绘制网络拓扑图；

2．配置路由器，包括基本配置、路由配置、GRE 配置等。通过查看路由表情况，验证 GRE 配置是否正确。

练习题

简答题

（1）说明 GRE 主要解决什么问题及其应用环境。

（2）描述 GRE 工作原理。

（3）说明 GRE 的配置命令及步骤。

项目 14　IPSec 隧道协议

在前面介绍的 GRE 隧道协议，可以在 LAN 之间通过 Internet 建立隧道，实现网络间资源共享，但是，GRE 隧道协议不能实现加密功能，传输数据不受加密保护。为了实现在隧道中传输数据包受到加密保护，需要使用 IPSec 隧道协议。

本章介绍 IPSec 隧道协议的定义、工作原理及相关配置命令。

知识点、技能点

1. 了解 IPSec 隧道协议的定义;
2. 掌握 IPSec 隧道协议的工作原理;
3. 掌握 IPSec 隧道协议的配置命令的使用方法。

14.1　问题提出

某公司网络由北京总公司及上海分公司构成。根据公司业务工作需要，总公司与分公司间建立 VPN 网络，传输公司专用数据，VPN 网络采用 IPSec 技术实现。

14.2　相关知识

VPN 是虚拟专用网络（Virtual Private Network）的缩写。VPN 提供一种通过公共网络对企业内部专用网络进行远程安全访问的连接方式。通过 VNP 技术，企业内部之间的数据流可以安全、透明地传输，从而提高网络系统的安全性和服务质量，如图 14.1 所示。

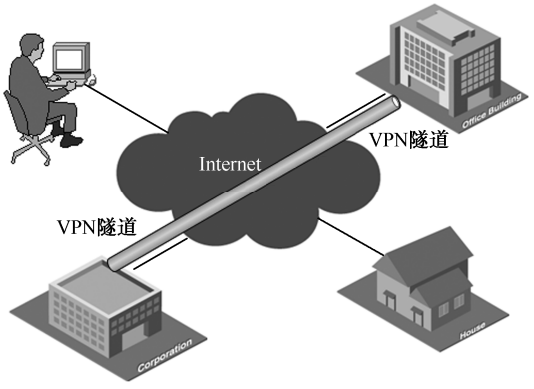


图 14.1　VPN 网络

VPN 提供的安全服务主要有以下三个方面。

第一方面，机密性服务：保证通过公共网络（如 Internet 等）传输的信息即使被他人截获也不会泄露。

第二方面，完整性服务：防止传输的数据被修改，保证数据的完整性、合理性。

第三方面，验证服务：提供用户和设备的访问身份验证，鉴别用户真实身份，防止非法接入。

有多种方式可以实现 VPN 功能。按照 VPN 在 TCP/IP 协议体系中的位置和实现方式不同，可将它分为链路层 VPN、网络层 VPN 和传输层 VPN。

目前，常见的 VPN 实现方案有：基于链路层的 VPN，如点对点隧道协议（PPTP）、二层转发协议（L2FP）、二层隧道协议（L2TP）等；基于网络层的 VPN，如 IP 安全协议（IPSec）；传输层 VPN，如安全套接字层协议（SSL）等。

实现 VPN 的关键技术是隧道，利用隧道协议实现数据封装。隧道协议有第二层隧道协议和第三层隧道协议。第二层隧道协议用于传输链路层协议；第三层隧道协议用于网络层协议。下面主要介绍实现网络层 VPN 的三层隧道协议 IPSec，其他协议读者可自行查阅相关资料。

1. IPSec 隧道协议

Internet 协议安全（IPSec）是一种开放标准的安全框架结构，是基于 AH、ESP 等安全协议基础上的网络层隧道协议。IPSec 能够在 Internet 协议 (IP) 网络上建立保密、安全的隧道，实现身份验证、数据完整性检查、保密等安全功能。

IPSec 在两台设备开始传递数据之前，它们之间通过协商，建立安全联盟 SA（Securtity Association），搭建传递数据的安全通道。所采用的协商方法可以是手工配置或采用标准的 IKE（Internet Key Exchange）协议方式。

在 IPSec 中有两个主要的安全协议：鉴别首部 AH（Authentication Header）协议和封装安全有效载荷 ESP（Encapsulation Security Payload）协议。AH 提供源点身份鉴别和数据完整性检查，但不提供保密。ESP 比 AH 复杂得多，它除提供源点身份鉴别及数据完整性检查外还提供保密功能。下面介绍几个主要术语。

（1）对等体（Peer）：是指参与 IPSec 的设备或其他设备，如网关路由器、计算机等。

（2）安全联盟：为特定数据流提供安全服务的一个具有某种约定的逻辑连接。这个逻辑连接中的约定参数包括特定的安全协议、安全算法、密钥及数据流描述。这里有 IPSec 安全联盟和 IKE 安全联盟两种，IPSec 安全联盟对数据提供 IPSec 保护功能，既可以由用户手工配置建立连接，又可以由 IKE 协商建立；IKE 安全联盟用于保护 IKE 的协商数据。

（3）安全参数索引（SPI）：SPI 是一个 32bit 的整数，和一个目的 IP 地址、一个安全协议类型一起，形成了安全联盟的唯一标志。当使用 IKE 来建立安全联盟时，每个安全联盟的 SPI 值都是一个唯一的随机数字。如果不使用 IKE，则需要手工为每个安全联盟指定 SPI 值。

（4）安全联盟生命期：安全联盟的有效期，手工方式建立起来的安全联盟没有生命期，也就是说可以永久使用，直到用户手工将其删除为止。IKE 协商的安全联盟的生命期是与对端 IKE 实体协商出来的，安全联盟生命一旦到期安全联盟就会被删除，IKE 重新协商新的安全联盟。

（5）变换集：变换集描述了由安全协议（AH 或 ESP）和算法组成的安全套件。例如，某变换集定义使用 ESP 协议和 DES 加密算法。

（6）加密映射条目：加密映射条目将变换集和数据流联系起来，并描述了对端的地址及通信必要参数，它完整地描述了与远端对等体的 IPSec 通信所需要的内容。通过加密映射条目，才能建立 IPSec 安全联盟。

2. IPSec 连接过程

以下描述 IPSec 的工作过程以通过 IKE 协商的工作过程为例。

（1）IPSec 的启动。当一个对等体向另一个对等体发送需要保护的数据流量时，则 IPSec 进程自动

启动。也可以通过手动启动。

（2）建立管理连接（IKE SA）。启动 IPSec 进程后，首先要做的事情就是建立 IKE 安全联盟，即建立管理连接。在这个阶段完成的主要任务如下。

任务 1：策略协商。确认采用什么样的加密方法（AES-256、AES-192、AES-128、3DES、DES 等）、采用什么样的完整性检查方法（SHA1、MD5 等）、采用什么样的验证方法（KerberosV5、证书或预共享密钥等）、采用什么样的创建密钥方法（DH Group14、DH Group2、DH Group1 等）。

任务 2：验证身份。设备之间通过相互交换对方身份信息验证对方身份。如共享秘钥验证等。

任务 3：交换参数并创建管理密钥。IKE 的一个主要特性就是不在网络上传递密钥，只在网络上传递创建密钥需要的相关参数，然后由通信双方计算机依据交换的相关参数分别产生相同管理密钥。管理密钥主要用于对管理连接中交换的数据加密和解密。

完成以上任务后，将建立一个安全的管理连接，用于设置之间交换管理信息。管理连接是一个双向连接，两台设备之间可以使用它来彼此共享 IPSec 消息。

（3）建立数据连接（IPSec SA）。完成 IKE 安全联盟建立后，可以创建 IPSec 数据连接。在这个阶段完成的主要任务如下。

任务 1：策略协商。确认 IPSec 采用什么安全协议（AH、ESP 等）、完整性与验证方法采用什么样的散列算法（MD5、SHA1 等）、采用什么样的加密方法（AES-256、AES-192、AES-128、3DES、DES 等）。

任务 2：创建会话密钥。可以使用阶段 1 交换的相关参数建立会话密钥，也可以重新交换另一组的相关参数，然后利用新的相关参数来创建会话密钥。会话密钥主要用于会话连接的用户数据加密和解密。

任务 3：将 SA、密钥及安全参数索引 SPI（Security Parameter Index）应用于驱动程序。SPI 是每对 SA 的标志符，每对 SA 有唯一的 SPI 值。

完成上述任务后，将建立一个安全的数据连接，用于传递用户数据信息。数据连接是一个单向连接，每个方向需要建立一个单独的数据连接。

（4）传输数据。当数据连接建立后，对等体间就可以通过这条数据连接通道安全地传输用户数据。

3. IPSec 配置

IPSec 配置是为了建立 IPSec 安全联盟。创建 IPSec 安全联盟有两种方式：手工配置 IPSec 安全联盟及通过 IKE 协商创建配置 IPSec 安全联盟。手工配置无须 IKE 介入，加密密钥由管理员手工设置，但需要指定更多的参数，安全性低；IKE 协商安全联盟中加密密钥由 IKE 产生，但需要对 IKE 参数进行配置，加密密钥可以定期自动协商更新，安全性较高。

IPSec 配置过程主要有以下任务。

- 创建加密访问列表：加密访问列表定义敏感数据范围，告诉设备要保护什么样的数据流。当满足有加密访问列表条件的数据流时启动 IPSec 保护功能。
- 定义变换集：变换集是安全协议、算法及封装模式的组合，告诉设备怎样保护敏感数据。当需要对数据进行保护时，对等体之间经协商采用一组相同的安全协议、算法及封装模式进行操作。
- 创建加密映射条目：创建加密映射条目就是将先前定义的访问列表和变换集合关联起来，并定义密钥，对等体地址，形成一套完整的 IPSec 解决方案。
- 将加密映射条目应用到接口上：这是一个激活定义 IPSec 解决方案的动作，将加密映射条目应用到接口上使加密映射条目在接口上开始工作。
- 配置默认生命周期（可选）：这是个可选步骤，可以通过这条命令来修改系统默认的生命周期

值。若无特别说明，IKE 将采用该生命周期值进行协商，从而使 IPSec 的生命周期将不会超过默认生命周期的长度。

■ 监视和维护 IPSec：监视和维护 IPSec，查看和调整 IPSec 参数，确定 IPSec 是否正常工作。

（1）创建加密访问列表（必须）。

加密访问列表用于定义哪些数据流要加密保护，哪些不要被加密保护，即解决对哪些数据实施加密保护的问题。IPSec 加密映射条目指定的加密访问列表有以下 4 个主要功能。

第一，对于发送路由器，明确对满足设定条件的数据流量是否进行 IPSec 加密保护，如使用 permit 则进行加密保护，如使用 deny 则不进行加密保护。

第二，对于发送路由器，明确 IPSec 安全联盟对什么样的数据流进行加密保护，使用加密访问列表的 permit 语句定义需要加密保护数据流量的范围。

第三，对于接收路由器，加密访问列表对于应该受到 IPSec 加密保护而没有保护的数据流量进行过滤和丢弃处理。

第四，在处理 IPSec 对等体发起的 IKE 协商时，决定是否接受代表所申请数据流的 IPSec 安全联盟申请（只有 ipsec-isakmp 加密映射条目需要协商）。必须确保两端对等体的访问列表匹配，我们建议，两端对等体访问列表配置成一致。

要配置加密访问列表，请在全局配置模式下执行以下命令：

ACL编号　　源地址

ruijie(config)# access-list *access-list-number* **{deny | permit}** *protocol source source-wildcard destination destination-wildcard*

设备名称为ruijie，处于全局配置模式　　协议　　目的地址

其中，*access-list-number* 为加密访问列表编号；*protocol* 为定义的加密访问列表使用的协议类型；*source* 为源地址；*source-wildcar*d 为源地址通配符；*destination* 为目标地址；*destination-wildacard* 为目标地址通配符。

加密访问列表使用数据流的通信协议、源地址、目的地址及其通配符来描述数据流。使用 permit 关键字将使得满足指定条件的所有 IP 通信都受到相应加密映射条目中所描述策略的加密保护。使用 deny 关键字可以防止通信受到特定加密映射条目的加密保护。

例如：创建加密访问列表保护子网 A（192.168.202.0/25）和子网 B（192.168.12.0/24）之间的所有 IP 通信。

```
ruijie(config)#access-list 120 permit ip 192.168.12.0 0.0.0.255 192.168.202.0 0.0.0.255
```

例如：创建加密访问列表保护主机 A（2.2.2.2）和主机 B（2.2.2.1）之间的 IP 通信。

```
ruijie(config)#access-list 101 permit ip host 2.2.2.2 host 2.2.2.1
```

例如：创建加密访问列表保护子网 A（192.168.12.0/24）和主机 C（202.101.11.3）之间的 TCP 通信。

```
ruijie(config)#access-list 120 permit tcp 192.168.12.0 0.0.0.255 202.101.11.3 0.0.0.0
```

例如：创建加密访问列表保护主机 D（1.1.1.1）和主机 E（2.2.2.2）之间的 Telnet 通信，其中主机 E 提供 Telnet 服务，则设备上定义：

```
ruijie(config)#access-list 133 permit tcp 1.1.1.1 0.0.0.0 2.2.2.2 0.0.0.0 eq telnet
```

我们建议，在本地对等体上定义的加密访问列表，都应该在远端对等体上定义镜像加密访问列表，否则，会导致一些数据得不到保护或者安全联盟无法建立。

（2）定义变换集（必须）。

变换集是 IPSec 所使用安全协议、算法以及封装模式的组合，解决对加密数据如何保护的问题。在 IPSec 安全联盟协商过程中，对等体必须使用同一个特定的变换集合来保护特定的数据流。

在加密映射条目中定义的变换集用于协商 IPSec 安全联盟，以保护由加密映射条目对应的访问列表所指定的那些数据流。在协商过程中，双方对等体寻找一个在双方对等体上都有的相同变换集。当找到了一个这样的变换集时，此变换集将被选中，并作为双方对等体 IPSec 安全联盟共同使用的安全参数运用到受保护的通信上。

通过手工来建立的安全联盟，由于手工建立的安全联盟不协商参数，因此在两端对等体上必须指定一个相同参数的变换集合。

如果对变换集的定义进行改变，那么改变对现存的安全联盟不会立即生效，只能应用于随后建立新安全联盟的协商中。如果想让这些新的设置马上生效，则可以通过使用 clear crypto sa 命令将安全联盟数据库清除。

定义变换集合的命令格式如下：

变换集合名称

ruijie(config)# crypto ipsec transform-set *transform-set-name transform1 [transform2 [transform3]]*

设备名称为ruijie
处于全局配置模式

变换集合组合

其中，*transform* 参数是系统所支持的变换集，常用的变换集（协议、算法组合）如表 14.1 所示。

表 14.1　常用变换集内容

变 换 集	说　明
ah-md5-hmac	AH 协议，MD5 HMAC 验证算法
ah-sha-hmac	AH 协议，SHA HMAC 验证算法
esp-des	ESP 协议，DES 加密算法
esp-3des	ESP 协议，3DES 加密算法
esp-des esp-md5-hmac	ESP 协议，采用 DES 加密算法和 MD5 HMAC 验证算法
esp-des esp-sha-hmac	ESP 协议，采用 DES 加密算法和 SHA HMAC 验证算法
esp-3des esp-sha	ESP 协议，采用 3DES 加密算法和 SHA HMAC 验证算法
esp-3des esp-md5	ESP 协议，采用 3DES 加密算法和 MD5 HMAC 验证算法
ah-md5-hmac esp-des	AH 协议在外，使用 MD5 HMAC 验证算法；ESP 协议在内，使用 3DES 加密算法

默认情况下，没有定义任何变换集。使用该命令的 no 形式删除一个变换集，如 **ruijie(config)# no crypto ipsec transform-set** *transform-set-name*。

例如：定义一个变换集名称为 myset，其中，保护方式为 ESP 协议，采用 DES 加密算法和 MD5 HMAC 验证算法。

```
ruijie(config)#crypto ipsec transform-set myset esp-des esp-md5-hmac
```

IPSec 提供两种封装模式，即传输模式（Transport）和隧道模式（Tunnel）。其中，传输模式实现在端到端的通信；而隧道模式实现站点到站点的通信。在变换集合配置模式下，可以修改 IPSec 的封装模式。修改封装模式的命令格式如下：

隧道模式

ruijie(cfg-crypto-trans)#mode {tunnel | transport}

设备名称为ruijie
处于变换集合配置模式

传输模式

其中，*tunnel* 为隧道模式；*transport* 为传输模式。默认情况下，系统为隧道模式（Tunnel）。使用该命令的 no 形式可以恢复默认模式，如 ruijie(cfg-crypto-trans)# no mode。

修改与变换集相关联的封装模式。封装模式设置只对那些源和目标地址都是 IPSec 对等体的通信有用；对于其他通信无用（所有其他通信都在通道模式下进行）。

例如：指定变换集的封装模式为通道模式。

```
ruijie(config)#crypto ipsec transform-set myset esp-des esp-md5-hmac
ruijie(cfg-crypto-trans)#mode tunnel
```

（3）创建加密映射条目（必须）。

在前面完成加密访问列表和变换集配置的基础上，还需要配置加密映射条目。加密映射条目将加密访问列表、变换集与设备关联起来，解决在哪个设备上对哪些数据实施什么加密保护的问题。配置加密映射条目需要完成以下几个方面的任务。

① 明确哪些数据流量应该受到 IPSec 加密保护，通过关联已配置好的加密访问列表来确定。

② 明确受到 IPSec 加密保护的数据流量将被发送到哪里，即远端 IPSec 对等体的 IP 地址是什么。通过命令指定远端对等体的 IP 地址。

③ 明确用于 IPSec 通信的本地地址是什么。将加密映射集合应用到接口上，IPSec 使用通信接口的地址作为本地对等体的地址。

④ 明确 IPSec 对于受保护的数据流量采取安全策略。从指定的一个或多个变换集中选择共同的安全协议、算法组合。

⑤ 安全联盟的生命周期是多少。如果没有特别设置，则采用默认的生命周期。

⑥ 安全联盟是被手工建立还是通过 IKE 建立的。通过命令指定。

加密映射条目同访问列表一样，也可以由多条语句构成，具有相同加密映射条目名称（但映射序列号不同）的加密映射条目组成一个加密映射集。将加密映射集应用到接口上，这样所有通过这个接口的 IP 数据流量都将被应用在接口上的加密映射集进行判断。

如果一个加密映射条目发现了应该受到保护的出站 IP 数据流量，并且加密映射指定了使用 IKE，将根据该加密映射条目中所包含的参数与远端对等体进行安全联盟协商；如果加密映射条目指定了使用手工建立的安全联盟，那么一个安全联盟在进行配置时就必须已经被建立好了。无论是手工建立还是 IKE 协商，只要安全联盟成功建立，数据将被加密传输，如果安全联盟协商失败，数据将被丢弃。

在加密映射条目中描述的策略将在安全联盟的协商过程中被使用，为了使两个 IPSec 对等体之间的 IPSec 创建工作能够顺利进行，两个对等体的加密映射条目必须包含互相兼容的配置语句。当两个对等体试图建立安全联盟时，双方都必须至少有一条加密映射条目和对方对等体的一条加密映射条目相兼容至少满足以下条件。

- 加密映射条目必须包含兼容的加密访问列表（如镜像映像访问列表）。
- 双方的加密映射条目都必须确定对方对等体地址（除非对等体正在使用动态加密映射）。
- 加密映射条目必须至少有一个相同的变换集。

对于单个接口只应用一个加密映射集。加密映射集中包含了 IPSec/IKE，或 IPSec/手工条目的组合。如果为给定的接口创建多个加密映射条目，那么就要使用映射条目的 seq- num 参数将这些映射条目排序；seq-num 值越小，优先级越高。

如果存在下面几种情况中的一种，就必须为一个接口创建多个加密映射条目。

- 如果此接口上不同的数据流将由不同的 IPSec 对等体进行处理。
- 如果想要对不同类型的通信（发往相同或不同的对等体）应用不同的 IPSec 安全性，例如，想让一组子网间的通信被认证，而另一组子网间的通信既被认证又被加密。在这种情况下，不同

类型的通信应该在两个不同的访问列表中被定义，并且必须为每个加密访问列表创建一个单独的加密映射条目。

加密映射集既可以用手工创建，也可以通过 IKE 创建，分别介绍如下。

方法一：手工创建安全联盟的加密映射集

第 1 步：创建加密映射集。

指定要创建加密映射集，执行此命令将进入加密映射集配置模式，命令格式如下：

加密映射条目名称

ruijie(config)#crypto map *map-name seq-num* **ipsec-manual**

设备名称为ruijie
处于全局配置模式

加密映射条目序号

其中，*map-name* 为加密映射集的名称；*seq-num* 为加密映射集的序列号；*ipsec-manual* 表示该映射集通过手工 IPSec 安全联盟。

默认情况下，没有配置任何加密映射集。该命令的 no 形式取消一个加密映射集，如 ruijie(config)# no crypto map map-name [seq-num]。

例如：手工创建一个 IPSec 安全联盟，名称为 mymap。

```
ruijie(config)#crypto map mymap 3 ipsec-manual
ruijie(config-crypto-map)#set peer 2.2.2.2
ruijie(config-crypto-map)#set session-key inbound esp 301 cipher abcdef1234567890
ruijie(config-crypto-map)#set sesession-key outbound esp 300 cipher abcdef1234567890
ruijie(config-crypto-map)#set transform-set myset
ruijie(config-crypto-map)#match address 101
```

第 2 步：指定受到 IPSec 加密保护的数据范围。

进入加密映射集配置模式后，使用访问列表定义哪些数据应该受到 IPSec 保护。这里指定的访问列表既用于出站通信也用于进站通信。在出站的情况下，检测到匹配的数据，如果已经有安全联盟，则对数据进行加密转发，如果没有建立安全联盟，将触发 IKE 创建安全联盟协商。在进站的情况下，检测到匹配的数据，如果是已经加密的数据，那么进行解密，如果是没有加密的数据，则直接丢弃。

指定加密映射集关联的访问列表命令如下：

访问控制列表编号

ruijie(config-crypto-map)# match address *access-list-number*

设备名称为ruijie
处于加密映射条目配置模式

其中，*access-list-number* 为访问列表编号（100～199，2000～2699，2900～3899），加密映射集只使用 IP 扩展访问列表。

默认情况下，加密映射集没有指定的访问列表。该命令的 no 形式删除加密映射集中的访问列表，如 ruijie(config-crypto-map)# no match address。

例如：由访问列表 101 定义名称为 mymap 加密映射集上的加密数据范围。

```
ruijie(config)#crypto map mymap 4 ipsec-isakmp
ruijie(config-crypto-map)#match address 101
```

第 3 步：指定远端对等体的 IP 地址。

对使用的加密映射集都必须指定一个远程对等体。指定加密映射集中远程对等体的命令格式如下：

对等体主机名称或IP

ruijie(config-crypto-map)# set peer *{hostname | ip-address }*

设备名称为ruijie
处于加密映射条目配置模式

其中，*hostname* 为远程对等体的主机名；*ip-address* 为远程对等体的 IP 地址。

默认情况下，不指定任何远程对等体。使用该命令的 no 形式删除加密映射集中的远程对等体，如 ruijie(config-crypto-map)# no set peer { hostname | ip-address }。

例如：在加密映射集 mymap 上指定一个远程对等体（2.2.2.2）。

```
ruijie(config)#crypto    map mymap 5 ipsec-isakmp
ruijie(config-crypto-map)#set peer 2.2.2.2
```

第 4 步：指定应用的变换集。

要为某个加密映射集指定使用哪些变换集。指定加密映射集使用的变换集命令格式如下：

变换集合组名称

ruijie(config-crypto-map)# set transform-set *transform-set-name1 [transform-set-name2] […] [transform-set-name6]*

设备名称为ruijie
处于加密映射条目配置模式

其中，*transform-set-name1* 为变换集 1 的名称。最多可以为加密映射集指定 6 个变换集。

默认情况下，没有指定任何变换集。使用该命令的 no 形式删除加密映射集和变换集的关联，如 ruijie(config-crypto-map)# no set transform-set。

例如：指定加密映射集的变换集为 myset。

```
ruijie(config)#crypto ipsec transform-set    myset    esp-des esp-sha-hmac
ruijie(config)#crypto    map mymap 5 ipsec-isakmp
ruijie(config-crypto-map)#set transform-set myset
```

第 5 步：指定安全参数索引（SPIs）和密码（手工配置 IPSec）。

上述指定的变换集中可以包含 AH 或 ESP 协议，AH 协议需要身份验证密码，ESP 需要身份验证密码及加密密码。另外，创建安全联盟需要使用安全参数索引（SPI）来标志一个具体的安全联盟，以便寻找合适的对等体建立 IPSec 安全联盟通信。指定安全参数索引和密码的命令对于 AH 协议和 ESP 协议，使用不同的命令格式。下面分别介绍手工配置 IPSec 时配置 AH 协议及 ESP 协议的安全参数索引及密码的命令格式。

对于 AH 协议下的指定安全参数索引及身份验证密码的配置命令格式如下：

入站或出站操作　　*16进制表示的密码值*

ruijie(config-crypto-map)# set session-key *{ inbound | outbound}* **ah** *spi hex-key-data*

设备名称为ruijie
处于加密映射条目配置模式　　*指定安全参数索引编号（SPI）*

其中，*inbound* 表示入站操作；*outbound* 表示出站操作；*spi* 为安全参数索引；*hex-key-data* 为一个 16 进制密码。

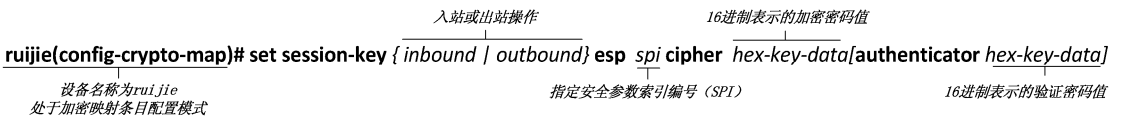
默认情况下，不指定任何安全参数索引和密码。该命令的 no 形式删除相关算法的安全参数索引和密码，如 ruijie(config-crypto-map)# no set session-key { inbound | outbound} ah。

上述指定安全参数索引和密码的命令只在 ipsec-manual 中使用。

例如：在加密映射集 mymap 中指定 AH 协议使用的身份验证密码。

```
ruijie(config)#crypto    map mymap 5 ipsec-manual
ruijie(config-crypto-map)#set session-key inbound ah 101 abcdef1234567890
ruijie(config-crypto-map)#set session-key outbound ah 100 abcdef1234567890
```

对于 ESP 协议下的指定安全参数索引及加密密码、身份验证密码的命令格式如下：



其中，*inbound* 表示入站操作；*outbound* 表示出站操作；*spi* 为安全参数索引；*cilpher hex-key-data* 为一个 16 进制加密密码；*authenticator hex-key-data* 为一个 16 进制验证密码。

如果变换集合包括了 ESP 加密算法（如 DES、3DES、AES、SM1 和 3DES 等）加密算法，必须给出加密密码（cilher 后面的 hex-key-data 内容）。如果变换集合包括了 ESP 验证算法（如 SHA、MD5 等），那么也必须给出验证密码。这里本地的进站 SPI 和密码要与远端对等体的出站 SPI、协议和密码要一致，反之亦然。

默认情况下，不指定任何安全参数索引和密码。该命令的 no 形式删除相关算法的安全参数索引和密码，如 ruijie(config-crypto-map)# no set session-key { inbound | outbound} esp。

上述指定安全参数索引和密码的命令只在 ipsec-manual 中使用。

说明：关键字 hex-key-data 是一个 16 进制，在具体配置的时候，只需输入 1～9、a～f 的字符串即可，不用加 0x 标志。如表 14.2 所示常用的算法长度。

表 14.2　常用算法的密码长度

算法名称	密码长度（bit）	输入的 16 进制串长（按字节）	配置实例
DES	64	8	set session-key inbound esp 300 cipher abcdef1234567890
3DES	192	24	set session-key inbound esp 300 cipher abcdef1234567890 abcdef1234567890abcdef1234567890
AES	128	16	set session-key inbound esp 300 cipher abcdef1234567890abcdef1234567890
SM1	128	16	set session-key inbound esp 300 cipher abcdef1234567890abcdef1234567890
MD5	128	16	set session-key inbound esp 302 cipher abcdef1234567890 authenticator abcdef1234567890abcdef1234567890
SHA	160	20	set session-key inbound esp 302 cipher abcdef1234567890 authenticator abcdef1234567890abcdef1234567890abcd

例如：为加密映射集 mymap 指定 ESP 使用的加密密码。

```
ruijie(config)#crypto map mymap 5 ipsec-manual
ruijie(config-crypto-map)#set session-key inbound esp 301 cipher abcdef1234567890
ruijie(config-crypto-map)#set session-key outbound esp 300    cipher abcdef1234567890
```

例如：某网络有路由器 A 和路由器 B，它们之间需要建立 IPSec 连接。

本地对等体（路由器 A）配置如下。

定义一个转换集 myset 及创建访问控制列表：

```
ruijie(config)#crypto ipsec transform-set myset esp-des
ruijie(config)#access-list 101 permit ip 192.168.12.0 0.0.0.255 192.168.202.0
0.0.0.255
```

定义一个手工映射集 mymap：

```
ruijie(config)#crypto map mymap 3 ipsec-manual
ruijie(config-crypto-map)#set peer 2.2.2.2
ruijie(config-crypto-map)#set session-key inbound esp 301 cipher abcdef1234567890
ruijie(config-crypto-map)#set session-key outbound esp 300 cipher abcdef1234567890
ruijie(config-crypto-map)#set transform-set myset
ruijie(config-crypto-map)#match address 101
ruijie(config-crypto-map)#
```

远程对等体（路由器 B）配置如下。

定义一个变换集 myset 及创建访问控制列表：

```
ruijie(config)#crypto ipsec transform-set myset    esp-des
ruijie(config)#access-list 101 permit ip 192.168.202.0 0.0.0.255 192.168.12.0
0.0.0.255
```

定义手工映射集 mymap：

```
ruijie(config)#crypto map mymap 3 ipsec-manual
ruijie(config-crypto-map)#set peer 2.2.2.1
ruijie(config-crypto-map)#set session-key inbound esp 300 cipher abcdef1234567890
ruijie(config-crypto-map)#set session-key outbound esp 301 cipher abcdef1234567890
ruijie(config-crypto-map)#set transform-set myset
ruijie(config-crypto-map)#match address 101
ruijie(config-crypto-map)#
```

方法二：使用 IKE 来创建安全联盟的加密映射条目

第 1 步：创建加密映射条目。

要配置使用 IKE 来建立安全联盟的加密映射条目，在全局配置模式下执行以下命令：

加密映射条目名称

ruijie(config)#crypto map *map-name* *seq-num* **ipsec-isakmp**

设备名称为ruijie
处于全局配置模式　　*加密映射条目序号*

其中，*map-name* 为加密映射条目的名称；*seq-num* 为加密映射条目的序列号；*ipsec-isakmp* 表示指定映射条目用于建立 IKE 协商的 IPSec 安全联盟。

ISAKMP 是 Internet 安全联盟与密码管理协议（Internet Security Association and Key Management Protocol）的缩写，为 IPSec 提供了一个秘钥管理框架，并为安全属性的协商提供了协议支持，不同的秘钥管理算法都可以集成到这个框架中。

默认情况下，没有配置任何加密映射条目。该命令的 no 形式取消一个加密映射或条目，如 ruijie(config)# no crypto map map-name [seq-num]。

例如：通过 IKE 协商创建安全联盟。

```
ruijie(config)#crypto map mymap 4 ipsec-isakmp
ruijie(config-crypto-map)#set peer 2.2.2.2
```

```
ruijie(config-crypto-map)#set transform-set myset
ruijie(config-crypto-map)#match address 101
```

第 2 步：指定受到 IPSec 加密保护的数据范围。

同手动创建 IPSec 安全联盟时的指定加密映射条目使用的访问列表命令。

第 3 步：指定远端对等体的 IP 地址。

同手动创建 IPSec 安全联盟时的指定远端对等体的 IP 地址命令。

第 4 步：指定应用的变换集。

同手动创建 IPSec 安全联盟时的指定应用的变换集命令。

在手动创建 IPSec 安全联盟时，需要手工设置安全参数索引及密码。通过 IKE 协商创建安全联盟时，不需要手工设置安全参数索引及密码，但是需要对 IKE 进行相关配置。具体配置 IKE 的命令及步骤可参考稍后的详细介绍。

（4）将加密映射集应用到接口上。

完成加密映射集配置后，需要将加密映射集应用于接口上才能发挥效用。将加密映射集应用在接口上的配置命令格式如下：

加密映射集名称

ruijie(config-if)#crypto map *map-name*

设备名称为ruijie
处于接口配置模式

其中，*map-name* 为加密映射集名称。默认情况下，没有任何加密映射集应用到接口上。使用该命令的 no 形式从一个接口上取消加密映射集的关联，如 ruijie(config-if)#no crypto map *[map-name]*。

一个接口只能同时应用一个加密映射集，加密映射集则能同时应用到多个接口上。在处理通过此接口的 IPSec 通信时，将使用此接口的 IP 地址作为本地设备地址。

例如：将名称为 mymap 的加密映射集应用于接口 S2/0 上。

```
ruijie(config)#interface serial 2/0
ruijie(config-if)#crypto map mymap
```

（5）配置默认生命周期（可选）。

IPSec 安全联盟加密通信是建立在使用共享密钥的基础上的。为了确保安全，安全联盟将在经过一定时间或通信量之后超时，重新协商，并使用新的共享密钥。当设备在进行安全联盟协商时，使用对等体提议的和本地设备配置的生命周期值中较小者作为新安全联盟的生命周期。

安全联盟生命周期有两种：按时间计算的生命周期和按通信量计算的生命周期。这两个生命周期中无论哪个先到期，安全联盟都将超时。如果修改了生命周期，此改变的生命周期只在新协商的安全联盟时才生效，不会影响现存安全联盟。如果想让新的生命周期马上生效，可以使用 clear crypto sa 命令清除现运行的安全联盟，重新建立新的安全联盟，则新的生命周期马上就生效了。

配置默认的生命周期命令格式如下：

超时时间（S）

ruijie(config)#crypto ipsec security-association lifetime {seconds *seconds* **|kilobytes** *kilobytes}*

超时通信量（KB）

设备名称为ruijie
处于全局配置模式

其中，*seconds* 为安全联盟超时时间（秒）。默认值为 3600 秒（1 小时）。可以配置 0，表示不启用时间超时功能。

kilobytes 为安全联盟的超时通信量（以千字节计）。默认值为 4,608,000KB（以每秒 10 兆字节的速度持续通信 1 小时）。可配置为 0，表示不启用字节超时功能。

该命令的 no 形式将生命周期恢复成默认值，如 ruijie(config)#no crypto ipsec security-association lifetime {seconds | kilobytes }。

例如：将 IPSec 安全联盟的时间周期设为 2500 秒，通信量生命周期设为 2304000KB（以 10MB 速率通信半个小时）。

```
ruijie(config)#crypto ipsec security-association lifetime second 2500
ruijie(config)#crypto ipsec security-association lifetime kilobytes 2304000
```

周期值越短，则破解密钥的攻击越难成功，因为攻击者用来进行分析的同一密钥加密的数据越少。但是，生命周期越短，用于建立新安全联盟的 CPU 处理时间就越多。手工建立的安全联盟是没有生命周期的。

生命周期的工作原理如下。

经过了一定的时间（由 seconds 关键字指定）或已通信了一定字节的数据通信量（由 kilobytes 关键字指定），这两件事情无论哪件先发生，安全联盟（及相应的密钥）都将超时。新的安全联盟在原有安全联盟的生命周期极限值到达以前就开始进行协商，以确保当原有安全联盟超时的时候，已经有一个新的安全联盟可以使用了。新安全联盟在 seconds 生命周期超时前 30 秒，或经由这条隧道的数据通信量距 kilobytes 生命周期还有 256KB 时开始进行协商（根据哪个先发生）。如果在一个安全联盟的整个生命周期中都没有通信经过这条隧道，那么当此安全联盟超时的时候不会进行新安全联盟的协商。相应地，新的安全联盟只有当 IPSec 看见应该受到保护的分组时才开始进行协商。

时间生命周期和通信量生命周期不能同时为零，否则不能成功协商。本地配置不检查，需要用户自己保证。

清除现有的安全联盟的命令格式如下：

ruijie#clear crypto ipsec sa

设备名称为ruijie处于特权模式

若使用上述命令清除现有的安全联盟，将引起所有现存的安全联盟立即中断；以后的安全联盟将使用新的生命周期，否则，所有现存的安全联盟将依据原来配置的生命周期超时。

例如：将清除所有安全联盟。

```
ruijie#clear crypto sa
```

（6）监视和维护 IPSec。

某些配置改变只能在协商后续安全联盟时生效。如果想让新设置立即生效，则必须删除现存的安全联盟，这样它们才会使用新设置重新建立。对于手工建立的安全联盟，必须删除和重新建立它们，否则改变永远不会生效。

① 删除（并重新发起）IPSec 安全联盟的命令格式如下：

ruijie#clear crypto sa

设备名称为ruijie处于特权模式

上述命令清空整个安全联盟数据库，这也将删除所有活动的安全线程。

对等体IP地址或主机名称

ruijie#clear crypto sa peer *{ ip-address | peer-name}*

设备名称为ruijie处于特权模式

上述命令清空有特定对等体地址的安全联盟。

映射条目名称

ruijie#clear crypto sa map *map-name*

设备名称为ruijie处于特权模式

上述命令清空特定加密映射集合的安全联盟。

对等体IP地址 *安全参数索引*

ruijie#clear crypto sa spi *destination-address {ah | esp}* *spi*

设备名称为ruijie处于特权模式 *安全协议*

上述命令清空特定对等体 IP 地址及 SPI 的安全联盟。

② 查看 IPSec 的配置信息的命令格式如下：

ruijie#show crypto ipsec transform-set

设备名称为ruijie处于特权模式

上述命令可以查看变换集合的配置信息。

例如：使用 show crypto ipsec transform-set 命令显示变换集合配置信息。

```
ruijie#show crypto ipsec transform-set
transform set myset3: { esp-des,}
      will negotiate = {Tunnel,}
```

映射条目名称

ruijie#show crypto map *[map-name]*

设备名称为ruijie处于特权模式

上述命令可以查看全部或指定的加密映射条目信息。

例如：显示全部的加密映射条目信息。

```
ruijie#show crypto map
 Crypto Map:"mymap1" 1 ipsec-isakmp, (Complete)
        Extended IP access list 100
        Security association lifetime: 0 kilobytes/120 seconds(id=2)
        PFS (Y/N): N
        Transform sets = { myset3,  }
         Interfaces using crypto map mymap1:
              GigabitEthernet 1/1/0
```

ruijie#show crypto ipsec sa

设备名称为ruijie处于特权模式

上述命令可以查看 IPSec 安全联盟信息。

例如：显示安全联盟信息。

```
ruijie#show crypto ipsec sa
Interface: GigabitEthernet 1/0/0
        Crypto map tag:mymap, local addr 2.2.2.3   //目前的加密映射集名称为 mymap，使用本地地址 2.2.2.3
        media mtu 1500
        ================================
        sub_map type:static, seqno:7, id=0
        local    ident (addr/mask/prot/port): (2.2.2.3/0.0.0.0/0/0)
        remote    ident (addr/mask/prot/port): (2.2.2.2/0.0.0.0/0/0)
        PERMIT //保护 2.2.2.3 和 2.2.2.2 之间的通信
                  #pkts encaps: 0, #pkts encrypt: 0, #pkts digest 0
                  #pkts decaps: 0, #pkts decrypt: 0, #pkts verify 0
        #send errors 0, #recv errors 0
        //统计数据，依次为：封装包数，加密包数，摘要包数，拆封包数，解密包数，验证包数，发送错误，接收错误

        Inbound esp sas: //进入包处理的安全联盟，协议为 ESP
            spi:0x79b8e4bb (2042160315) //SPI 的值为 2042160315
            transform: esp-3des //变换集合为 ESP-3DES
            in use settings={Tunnel,} //通道模式
            crypto map mymap 7
            sa timing: remaining key lifetime (k/sec): (4607000/3505)
            //离安全联盟的生命周期到期还有：4607000 千字节/3505 秒
            IV size: 8 bytes       //IV 向量长度为 8
        max reply windows size: 0
        Replay detection support:Y //抗重播处理

       Outbound esp sas:   //外出包处理的安全联盟，协议为 ESP
          spi:0x293b8b55 (691768149) //SPI 的值为 691768149
          transform: esp-3des    //变换集合为 ESP-3DES
          in use settings={Tunnel,} //通道模式
          crypto map mymap 7
          sa timing: remaining key lifetime (k/sec): (4607000/3505)
          //离安全联盟的生命周期到期还有：4607000 千字节/3505 秒
          IV size: 8 bytes      //IV 向量长度为 8
          max reply windows size: 0
          Replay detection support:Y    //抗重播处理
```

4. IKE 配置

IKE 是 Internet 秘钥交换协议（Internet Key Exchange Protocol）的缩写，与 IPSec 一起使用。当 IPSec

配置为通过 IKE 创建安全联盟时，在接口上检测到有符合要求的外出数据包时，IPSec 触发本地 IKE 和远端对等体 IKE 进行协商，在对等体之间建立一个 IKE 安全联盟，通过 IKE 联盟传递所支持的各种 IPSec 参数，最终在两端创建一致的 IPSec 安全联盟，实现对外出数据包的保护。经过一段时间，IPSec 安全联盟的生命期到达时，如果还有符合要求的数据需要传送，双方的 IKE 将重新开始协商 IPSec，这样周而复始。

使用 IKE 避免了在通信双方的加密映射表中手工指定所有的 IPSec 参数和密钥；能指定 IPSec 安全联盟的生命期限；使 IPSec 定期更换密钥，增加安全性；IKE 使 IPSec 能提供抗重播服务。

IKE 配置任务包括以下几个方面。

第一，启动 IKE：使用时，确保 IKE 正在工作，而没有被关闭。

第二，确保访问列表与 IKE 兼容：如果设备配置了访问列表，确保没有将 IKE 的通信流禁止。

第三，选择工作模式：IKE 协商的时候存在两种工作模式（主模式和积极模式）（注：积极模式在某些文献上也称为野蛮模式）。

第四，创建 IKE 策略：指定 IKE 协议使用的参数。

第五，配置预共享密钥：IKE 和对等体之间共同的密钥。

第六，IKE 维护（可选）：对 IKE 进行维护，查看参数，确认 IKE 正常工作。

（1）开启 IKE。

默认情况下，IKE 功能是打开的。如果不想将 IKE 与 IPSec 实现一起使用，可以使用命令关闭它，但此时只有手工方式的 IPSec 安全联盟才能工作。

要启动 IKE，执行以下命令：

ruijie(config)#crypto isakmp enable

设备名称为ruijie
处于全局配置模式

使用该命令的 no 形式禁止 IKE，如 ruijie(config)# no crypto isakmp enable。

例如：使能 IKE。

```
ruijie(config)#crypto isakmp enable
```

（2）确保 IKE 流量通过。

IKE 是运行在 UDP 上的应用程序，其通信报文为 UDP 数据，接口为 500。如果在设备上配置了访问列表（防火墙），并将此通信报文禁止，将导致 IKE 协商失败。所以，必须确保 IKE 的通信报文没有被禁止。

（3）选择工作模式。

IKE 协商的过程中分为以下两个阶段。

第一阶段：实现在两个 ISAKMP 实体之间建立安全的、经过验证的通道来进行通信，该阶段采用的模式有主模式和积极模式。

第二阶段：主要是协商代表服务的安全联盟（SA）。

在第一阶段中有两种工作模式，根据各自的优缺点，在默认的情况下，我们采用主模式，但在 IP 地址不固定的情况下，可以使用积极模式。

配置 IKE 工作模式的命令格式如下：

工作模式选项

ruijie(config-crypto-map)# set exchange-mode *{main | aggressive}*

设备名称为ruijie
处于加密映射条目配置模式

其中，*main* 为主模式；*aggressive* 为积极模式。使用该命令的 no 形式可以恢复为默认模式，如 ruijie(config-crypto-map)#no set exchange-mode。

例如：设置 IKE 的工作模式为积极模式。

```
ruijie(config)#crypto map mymap 10    ipsec-isakmp
ruijie(config-crypto-map)#set exchange-mode aggressive
```

（4）创建 IKE 策略。

参与 IKE 协商的双方至少拥有一套一致的 IKE 策略，这是 IKE 协商成功的必要条件。必须在每台对等体上创建多个具有优先级的策略，以确保至少有一个策略能够匹配远程对等体的策略。在每个 IKE 策略中定义如表 14.3 所示的 5 个参数。

表 14.3　IKE 策略参数表

参　数	关 键 字	可 取 值	默 认 值
加密算法	des	56 比特 DES-CBC	56 比特 DES-CBC
	3des	168 比特 3DES-CBC	
HASH 算法	sha	SHA-1（HMAC 变体）	SHA-1（HMAC 变体）
	md5	MD5（HMAC 变体）	
验证方法	pre-share	预共享密钥	数字签名验证
	rsa-sig	数字签名验证	
Diffie-Hellman 组标志	1	768 比特 Diffie-Hellman 组	768 比特 Diffie-Hellman 组
	2	1024 比特 Diffie-Hellman 组	
IKE 安全联盟的生命周期	空	1 分钟到一天内的秒数	一天（86400 秒）

IKE 协商开始时，IKE 试图去寻找在两个对等体上一致的策略。发起协商的一方将策略全部发送给远程响应方，响应方在接收到的另一端对等体的策略中按优先顺序寻找与本地匹配的策略。

当协商双方的策略均包括相同的加密算法、HASH 算法、验证方法及 Diffie-Hellman 参数值，并且当远程对等体的策略所指定的生命周期小于或等于所比较的策略生命周期时作出匹配选择（如果没有指明生命周期，则将使用较短的远程对等体的策略生命周期）。如果没有发现可接受的匹配策略，IKE 拒绝协商，从而不会建立 IPSec。如果找到一种匹配的策略，IKE 将完成协商，并且创建 IPSec 安全联盟。

可以在安全性和性能之间进行平衡，选择需要协商的每个参数值。

- 加密算法：当前支持 56 比特 DES-CBC，168 比特 3DES-CBC，128 比特 AES-CBC，128 比特 SM1-CBC。
- HASH 算法：SHA-1 和 MD5。MD5 的摘要信息较少，通常认为要比 SHA-1 稍稍快一些，SHA-1 安全性更高些。
- 验证方法：RGOS 目前支持预共享方式验证和数字签名验证。采用预共享方式验证双方需要配置正确的预共享密钥，如果采用数字签名验证双方需要正确的证书配置（自行参阅相关资料了解证书配置过程及命令格式）。

- Diffie-Hellman 组标志有两个选项：768 比特或 1024 比特 Diffie-Hellman。1024 比特 Diffie-Hellman 较难以攻克，但会占用更多的 CPU 时间。
- IKE 安全联盟的生命周期不同于 IPSec 的安全联盟的生命周期，它指的是 IKE 参数协商的有效期，可以设置为任意值。作为一条通用的规则，生命周期越短，IKE 协商将越安全。

配置 IKE 策略的命令格式如下。

第 1 步：创建 IKE 策略。为了配置 IKE 策略中的相关参数，需要创建一条具有一定优先级的策略语句，并进入 IKE 策略配置模式。创建 IKE 策略的命令格式如下：

IKE策略的优先级

ruijie(config)#crypto isakmp policy *priority*

设备名称为ruijie
处于全局配置模式

其中，*priority* 为 IKE 策略的优先级。使用 1～10000 的整数，1 是最高优先级而 10000 是最低优先级。默认情况下，没有默认优先级。使用该命令的 no 形式删除某个优先级的策略，如 ruijie(config)# no crypto isakmp policy priority。

可以创建多重 IKE 策略，每条策略对应不同的参数值组合（加密、HASH、验证及 Diffie-Hellman 参数值等）。对所创建的每一条策略，需分配唯一的优先权，从优先级高的开始检查匹配情况。如果不配置任何策略，则设备使用默认的策略，该策略被设为最低优先级，并包含每个参数的默认值。

例如：配置一个优先级为 100 的 IKE 策略。

```
ruijie(config)#crypto isakmp policy 100
ruijie(isakmp-policy)#authentication pre-share
ruijie(isakmp-policy)#encryption des
ruijie(isakmp-policy)#group   2
ruijie(isakmp-policy)#hash   sha
ruijie(isakmp-policy)#
```

第 2 步：指定加密算法。IKE 安全联盟需要指定对联盟传输数据进行加密的加密算法是什么。该加密算法不同于 IPSec 安全联盟的加密算法。指定 IKE 安全联盟加密算法的命令格式如下：

IKE策略的加密算法

ruijie(isakmp-policy)#encryption *{des|3des|aes-128|aes-192|aes-256|sm1}*

设备名称为ruijie
处于IKE策略配置模式

其中，*des* 指定 56 比特的 DES-CBC 作为加密算法；*3des* 指定 168 比特的 3DES-CBC 作为加密算法；*aes-128* 指定密钥长度为 128 比特的 AES 作为加密算法；*aes-192* 指定密钥长度为 192 比特的 AES 作为加密算法；*aes-256* 指定密钥长度为 256 比特的 AES 作为加密算法；*sm1* 指定密钥长度为 128 比特的 SM1 作为加密算法。

默认情况下，指定 56 比特的 DES-CBC 加密算法。使用该命令的 no 形式恢复默认值，如 ruijie(isakmp-policy)# no encryption。

例如：指定 IKE 策略中的加密算法为 DES。

```
ruijie(config)#crypto   isakmp policy 10
ruijie(isakmp-policy)#encryption des
```

第 3 步：指定 HASH 算法。指定 IKE 策略中的 HASH 算法的配置命令格式如下：

IKE策略的HASH算法

ruijie(isakmp-policy)#hash *{sha | md5}*

设备名称为ruijie
处于IKE策略配置模式

其中，*sha* 指定 SHA-1（HMAC 变体）作为 HASH 算法；*md5* 指定 MD5（HMAC 变体）作为 HASH 算法。默认情况下，HASH 算法为 SHA。使用该命令的 no 形式恢复 HASH 算法为默认值，如 ruijie(isakmp-policy)# no hash。

例如：指定 IKE 策略中的 HASH 算法为 MD5。

```
ruijie(config)#crypto　isakmp policy 10
ruijie(isakmp-policy)#hash md5
```

第 4 步：指定验证方法。目前，IKE 协商策略中验证方式默认采用数字签名验证，这与 Cisco 相同，如果希望采用预共享密钥方式验证则需要增加一个 IKE 策略（配置成预共享方式）。digital-email 验证模式仅用于在特殊场合下和通过 SM1 算法验证的设备互通。发起端需要预先配置对端证书。指定 IKE 策略的验证方法的命令格式如下：

IKE策略的验证方法

ruijie(isakmp-policy)#authentication *{pre-share|rsa-sig|digital-email }*

设备名称为ruijie
处于IKE策略配置模式

其中，*pre-share* 为预共享密钥验证；*rsa-sig* 为数字签名验证；*digital-email* 为数字信封验证（来自《IPSec VPN 技术规范》）。默认情况下，RGOS 8.31 版本以后为数字签名验证，RGOS 8.31 版本前为预共享密钥验证。使用该命令的 no 形式恢复默认验证方法，如 ruijie(isakmp-policy)#no authentication。

例如：指定 IKE 策略中的验证方法为预共享验证。

```
ruijie(config)#crypto　isakmp policy 10
ruijie(isakmp-policy)#authentication pre-share
```

第 5 步：指定 Diffie-Hellman 组标志。指定 IKE 策略中的 Diffie-Hellman 组标志的命令格式如下：

DH组标志

ruijie(isakmp-policy)#group *{1|2}*

设备名称为ruijie
处于IKE策略配置模式

其中，“1”表示指定 768 比特 Diffie-Hellman 组；“2”表示指定 1024 比特 Diffie-Hellman 组。默认情况下，指定 768 比特 Diffie-Hellman 组（group 1）。使用该命令的 no 形式恢复 Diffie-Hellman 组标志为默认值，如 ruijie(isakmp-policy)# no group。

例如：指定 IKE 策略的 Diffie-Hellman 组为 1024 比特。

```
ruijie(config)#crypto　isakmp policy 10
ruijie(isakmp-policy)#group 2
```

第 6 步：指定 IKE 安全联盟的生命周期。当 IKE 开始协商时，它所做的第一件事情是为其对话在安全参数上达成一致。这些一致的参数就在每个对等体上由 IKE 安全联盟（SA）引用，在每个对等体上保留直到 IKE 安全联盟生命期超时。配置 IKE 安全联盟生命周期的命令格式如下：

IKE生命周期值（秒）

ruijie(isakmp-policy)#lifetime *seconds*

设备名称为ruijie
处于IKE策略配置模式

其中，*seconds* 为 IKE 生命周期值（秒）。使用 60～86400 秒之间的整数。默认情况下，IKE 生命周期值为 86400 秒（1 天）。使用该命令的 no 形式恢复生命周期为默认值，如 ruijie(isakmp-policy)# no lifetime。

例如：设置 IKE 安全联盟的生命周期为 1000 秒。

```
ruijie(config)#crypto    isakmp policy 10
ruijie(isakmp-policy)#lifetime 1000
```

（5）配置预共享密钥。

预共享密钥是参与 IKE 协商的两个对等体之间共同拥有的密钥，因此每个预共享密钥对应一对 IKE 对等体。为了安全起见，在不同的对等体对之间配置不同的密钥。

要配置预共享密钥，请在全局配置模式下执行以下命令：

明文或密文　　*远端主机名称*

ruijie(config)#crypto isakmp key *{ 0 | 7 }* *keystring* **{ hostname** *peer-hostname* **|address** *peer-address [mask]* **}**

设备名称为ruijie
处于全局配置模式　　*密钥字符串*　　*远端主机IP地址及掩码*

其中，“*0*”表示密钥显示为明文；“*7*”表示密钥显示为密文；*keystring* 为预共享密钥字符串，最长可以使用 128 个字符；*peer-hostname* 为远程对等体的主机名；*peer-address* 为远程对等体的 IP 地址；*mask* 为指定的 IP 地址为一个网段的地址时的子网掩码。

默认情况下，没有指定任何共享密钥。使用该命令的 no 形式删除指定的预共享密钥，如 ruijie(config)#no crypto isakmp key { 0 | 7} keystring { hostname peer-hostname |address peer-address [mask] }。

如果采用 hostname 来标志对端身份，需要使用如下命令指定该 hostname 所对应的 IP 地址。

ruijie(config)#ip host *hostname address*

IKE 一般使用预共享协商，要使 IKE 成功建立 IKE 安全联盟，必须使用该命令在进行通信的两个对等体上配置相同的预共享密钥。如果指定的对等体为一个网段，采用 mask 来标志它的子网掩码。

例如：指定与对等体 172.16.1.1 进行 IKE 协商的预共享密钥为 mysecret。

```
ruijie(config)#crypto isakmp key 0    mysecret address 172.16.1.1
```

（6）IKE 维护。

显示 IKE 连接相关信息的命令格式如下：

ruijie#show crypto isakmp sa

设备名称为ruijie处于特权模式

例如：显示 IKE 联盟信息。

```
ruijie#show crypto    isakmp sa
  destination     source      state       conn-id     lifetime(second)
  1.1.1.1         1.1.1.2     IKE_IDLE    59          32254
```

显示 IKE 策略参数的命令格式如下：

ruijie#show crypto isakmp policy

设备名称为ruijie处于特权模式

例如：显示 IKE 策略参数信息。

```
ruijie#show crypto isakmp policy
Protection suite of priority 9
encryption algorithm:    3DES - Data Encryption Standard (56 bit keys)
hash algorithm:          Message Digest 5
authentication method:   Pre-Shared Key
Diffie-Hellman group:    #2 (1024 bit)
lifetime:                1000 seconds
Protection suite of priority 10
encryption algorithm:    DES - Data Encryption Standard (56 bit keys)
hash algorithm:          Message Digest 5
authentication method:   Pre-Shared Key
Diffie-Hellman group:    #2 (1024 bit)
lifetime:                1000 seconds
Default protection suite
encryption algorithm:    DES - Data Encryption Standard (56 bit keys)
hash algorithm:          Secure Hash Standard
authentication method:   Rsa-Sig
Diffie-Hellman group:    #1 (768 bit)
lifetime:                86400seconds
```

清除 IKE 连接的命令格式如下：

ruijie#clear crypto isakmp *[connection-id]*

设备名称为ruijie处于特权模式

其中，*connection-id* 为 IKE 安全联盟的 ID 号。如果没有使用 connection-id 参数，则该命令清除所有存在的 IKE 安全联盟。

5. 配置实例

某公司由总公司与分公司构成，总公司网络地址为 202.126.101.0/24，分公司网络地址为 67.151.69.0/24。总公司与分公司通过 Internet 建立 IPSec 连接，保护两个子网之间的 IP 数据通信，R1 路由器作为子网 202.126.101.0 的网关，R2 路由器作为子网 67.151.69.0 的网关，如图 14.2 所示。要求实现以下目的：

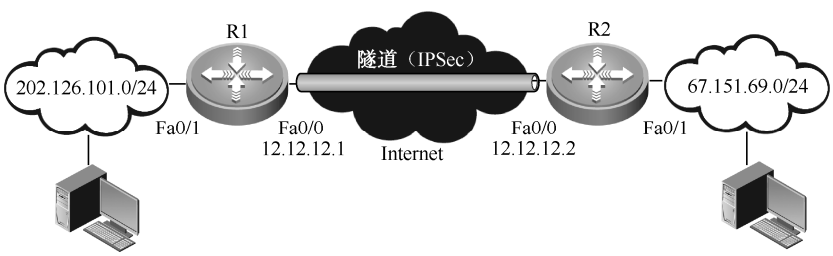


图 14.2　IPSec 应用（一）

- 阶段一协商采用 3DES 算法加密。
- 采用通道模式。
- 保护方式为 ESP-DES-MD5（提供加密和验证服务）。

路由器 R1 的参数配置如下。

（1）配置 IKE 安全联盟。

第 1 步：开启 IKE。

```
R1(config)#crypto isakmp enable
```

第 2 步：配置 IKE 策略。

```
R1(config)#crypto isakmp policy 1
R1(config-isakmp)#authentication pre-share
R1(config-isakmp)#encryption 3des
R1(config-isakmp)#exit
```

第 3 步：配置 IKE 预共享密钥。

```
R1(config)#crypto isakmp key preword address 12.12.12.2
```

（2）配置 IPSec 安全联盟。

第 1 步：配置变换集合。

```
R1(config)#crypto ipsec transform-set myset    esp-des esp-md5-hmac
```

第 2 步：定义一个加密映射条目。

```
R1(config)#crypto map mymap    5    ipsec-isakmp
R1(config-crypto-map)#set peer 12.12.12.2
R1(config-crypto-map)#set transform-set myset
R1(config-crypto-map)#match address 101
R1(config-crypto-map)#exit
```

（3）配置接口 IP 地址及应用映射条目。

```
R1(config)#interface FastEthernet 0/0
R1(config-if)#ip address 12.12.12.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#crypto map mymap
R1(config-if)#exit
R1(config)#interface FastEthernet 0/1
R1(config-if)#ip address 202.126.101.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
```

（4）定义加密访问列表。

```
R1(config)#access-list 101 permit ip 202.126.101.0 0.0.0.255    67.151.69.0 0.0.0.255
```

（5）设置默认路由。

```
R1(config)#ip route 0.0.0.0 0.0.0.0 Fa 0/0
R1(config)#end
```

路由器 R2 的参数配置如下。

（1）配置 IKE 安全联盟。

第 1 步：开启 IKE。

```
R2(config)#crypto isakmp enable
```

第 2 步：配置 IKE 策略。

```
R2(config)#crypto isakmp policy 1
R2(config-isakmp)#authentication pre-share
R2(config-isakmp)#encryption 3des
R2(config-isakmp)#exit
```

第 3 步：配置 IKE 预共享密钥。

```
R2(config)#crypto isakmp key preword address 12.12.12.1
```

（2）配置 IPSec 安全联盟。

第 1 步：配置变换集合。

```
R2(config)#crypto ipsec transform-set myset    esp-des esp-md5-hmac
```

第 2 步：定义一个加密映射条目。

```
R2(config)#crypto map mymap    5    ipsec-isakmp
R2(config-crypto-map)#set peer 12.12.12.1
R2(config-crypto-map)#set transform-set    myset
R2(config-crypto-map)#match address 101
R2(config-crypto-map)#exit
```

（3）配置接口 IP 地址及应用映射条目。

```
R2(config)#interface FastEthernet 0/0
R2(config-if)#ip address 12.12.12.2 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#crypto map mymap
R2(config-if)#exit
R2(config)#interface FastEthernet 0/1
R2(config-if)#ip address 67.151.69.1 255.255.255.0
R2(config-if)#no shutdown
R2(config-if)#exit
```

（4）定义加密访问列表。

```
R2(config)#access-list 101 permit ip 67.151.69.0 0.0.0.255 202.126.101.0 0.0.0.255
```

（5）设置默认路由。

```
R2(config)#ip route 0.0.0.0 0.0.0.0 Fa 0/0
R2(config)#end
R2#
```

14.3　扩展知识

1. AH 协议

AH 是验证头协议（Authentication Header）的缩写。它为 IP 通信提供数据源验证、数据完整性和抗重播保证，它能保护通信免受篡改，但不能防止窃听，适合用于传输非机密数据。AH 的工作原理是在每一个数据包上添加一个身份验证报头。此报头包含一个带密钥的 HASH 散列（可以将其当作数字签名，只是它不使用证书），此 HASH 散列在整个数据包中计算，因此对数据的任何更改将致使散列无效，这样就提供了完整性保护。

（1）AH 协议报头。

AH 协议为了实现身份验证、数据完整性和抗重播功能，依据被保护的 IP 数据包，构建一个 AH 协议报头。AH 协议报头结构如图 14.3 所示。

图 14.3　AH 协议报头

其中，下一个报头（Next Header）：识别下一个使用 IP 协议号的报头，例如，Next Header 值等于"6"，表示紧接其后的是 TCP 报头。

长度（Length）：AH 报头长度。

安全参数索引（Security Parameters Index，SPI）：安全联盟的唯一标志号，32 位伪随机值。SPI 值为 0 被保留，表示"没有创建安全联盟"。

序列号（Sequence Number）：从 1 开始的 32 位单增序列号，不允许重复，唯一地标志了每一个发送数据包，为安全联盟提供抗重播保护。接收端校验序列号为该字段值的数据包是否已经被接收过，若是，则拒收该数据包。

验证数据（Authentication Data）：包含完整性检验和。接收端接收数据包后，首先执行 HASH 计算，再与发送端所计算的该字段值比较，若两者相等，表示数据完整，若在传输过程中数据遭修改，两个计算结果不一致，则丢弃该数据包。

（2）传输模式下的 AH 协议。

AH 协议可以工作在两种模式中：传输模式和隧道模式。

在传输模式下，AH 协议对 IP 数据包的数据部分及 IP 报头固定字段内容进行 MD5 或 SHA-1 完整性计算，产生验证数据（HASH 检验和），也可以说 AH 协议在传输模式下，只对数据包及 IP 报头固定字段内容进行完整性保护。在传输模式下 AH 协议将 AH 报头添加到原 IP 报头与 IP 数据包数据部分之间，AH 协议数据结构如图 14.4 所示。

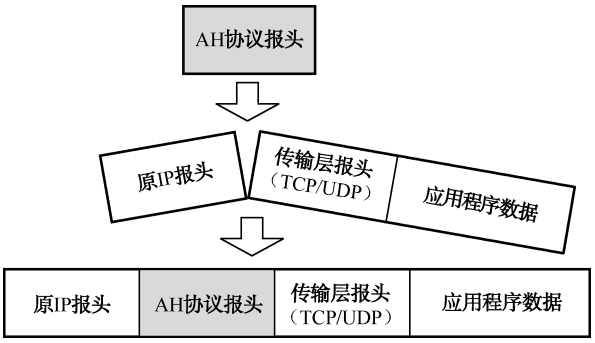


图 14.4　传输模式下的 AH 数据结构

可以看出，在传输模式下，IP 数据包的首部 IP 地址没有变化。传输模式适合于端到端的数据传输，如图 14.5 所示。

（3）隧道模式下的 AH 协议。

在隧道模式下，AH 协议对整个 IP 数据包进行 MD5 或 SHA-1 完整性计算，产生验证数据（HASH 检验和）。也可以说，AH 协议在隧道模式下，只对整个 IP 数据包进行完整性保护。在隧道模式下 AH

协议将 AH 报头封装在 IP 数据包前面，然后重新再封装一个新的 IP 报头，AH 协议数据结构如图 14.6 所示。

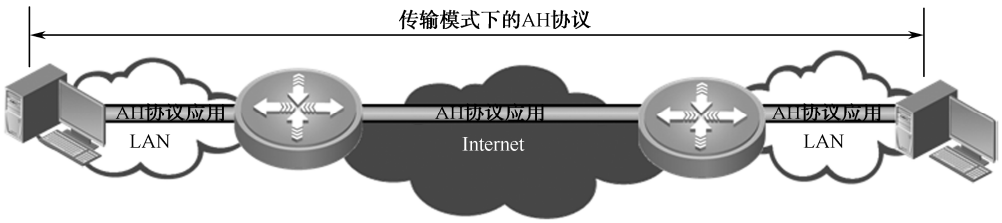


图 14.5　传输模式下的 AH 协议应用

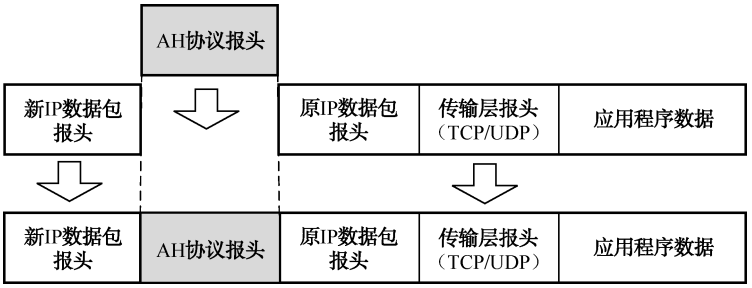


图 14.6　隧道模式下的 AH 数据结构

可以看出，在隧道模式下，原 IP 数据包的首部 IP 地址已经被封装起来了，新 IP 数据包的首部 IP 地址是在 AH 协议报头之前增加的。隧道模式下，一方面对整个数据包进行完整性保护，另一方面又新增加了一个 IP 首部，改变了 IP 地址。隧道模式适合于站点到站点的数据传输，如图 14.7 所示。

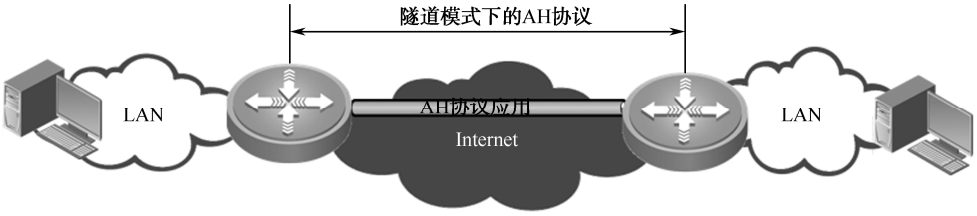


图 14.7　隧道模式下的 AH 协议应用

2. ESP 协议

ESP 是封装安全载荷协议（Encapsulated Security Payload）的缩写。它不但可以提供 IP 通信的数据源验证、数据完整性和抗重播保证，还可以为所传输的数据提供加密保护。它能保护通信免受篡改和泄露。

（1）ESP 协议报头与报尾。

ESP 协议与 AH 协议结构略有不同，ESP 协议结构分为 ESP 协议报头、报尾及验证报尾三部分。ESP 协议报头结构如图 14.8 所示。

图 14.8　ESP 协议报头

其中，安全参数索引（SPI）：安全联盟的唯一标志号，32 位伪随机值。SPI 值为 0 被保留，表示“没有创建安全联盟”。

序列号：从 1 开始的 32 位单增序列号，不允许重复，唯一地标志了每一个发送数据包，为安全联盟提供抗重播保护。接收端校验序列号为该字段值的数据包是否已经被接收过，若是，则拒收该数据包。

ESP 协议除报头外还有 ESP 报尾和 ESP 验证报尾两个结构，ESP 报尾和 ESP 验证报尾结构如图 14.9 所示。

图 14.9　ESP 协议报尾及验证报尾

其中，扩展位（Paading）：用于当加密算法要求明文长度为某一长度整数倍时，则可以通过填充达到所需长度等，范围为 0～255 字节。

扩展位长度（Pad Length）：指明扩展位的长度，接收端利用它恢复负载数据真实的长度，范围为 8bit。

下一个报头：识别下一个使用 IP 协议号的报头，例如，Next Header 值等于"6"，表示紧接其后的是 TCP 报头。

验证数据：包含完整性检验和。接收端接收数据包后，首先执行 HASH 计算，再与发送端所计算的该字段值比较，若两者相等，表示数据完整，若在传输过程中数据遭修改，两个计算结果不一致，则丢弃该数据包。

（2）传输模式下的 ESP 协议。

ESP 协议与 AH 协议一样，也可以工作在传输模式和隧道模式。

在传输模式下，ESP 协议将 ESP 协议报头加入到 IP 数据包报头与上层数据之间的位置，将 ESP 协议报尾及 ESP 协议验证报尾添加到上层数据之后位置，如图 14.10 所示。

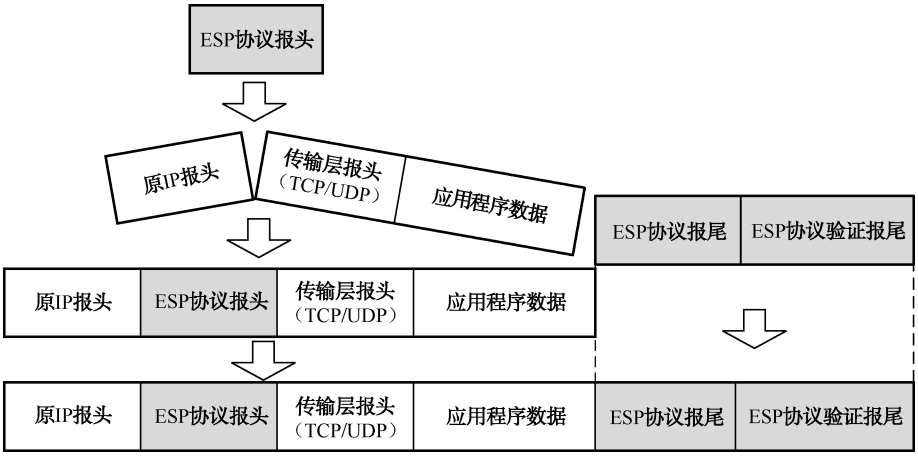


图 14.10　传输模式下 ESP 协议数据结构

ESP 协议除对传输的数据实现身份验证、完整性验证、防止重播外，还提供数据加密功能。在传输模式下，ESP 协议只是对 IP 数据包的数据部分及 ESP 协议报尾部分进行加密处理，如图 14.11 所示。

图 14.11　传输模式下 ESP 协议加密内容

传输模式下的 ESP 协议的应用如图 14.12 所示。

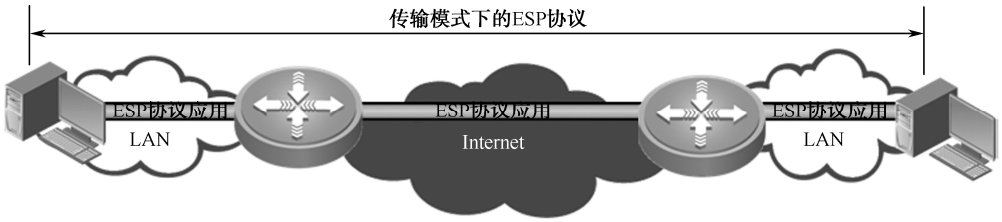


图 14.12　传输模式下 ESP 协议应用

（3）隧道模式下的 ESP 协议。

在隧道模式下，ESP 协议对整个 IP 数据包进行 MD5 或 SHA-1 完整性计算，产生验证数据（HASH 检验和）。也可以说，ESP 协议在隧道模式下，只对整个 IP 数据包进行完整性保护。在隧道模式下 ESP 协议将 ESP 报头封装在 IP 数据包前面，然后重新再封装一个新的 IP 报头，并且将 ESP 协议报尾及验证报尾添加到数据包后面，ESP 协议数据结构如图 14.13 所示。

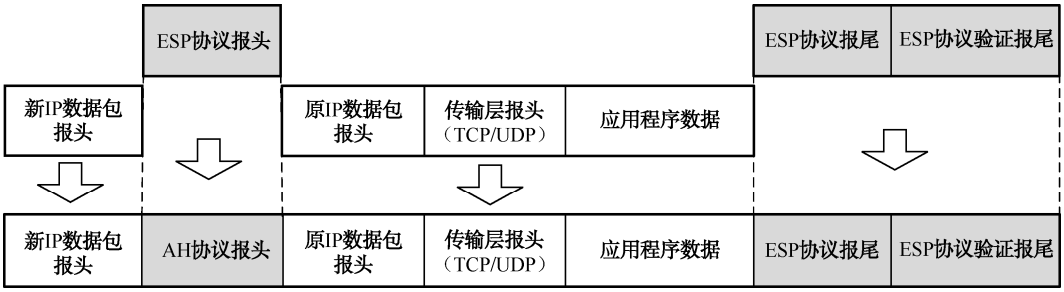


图 14.13　隧道模式下 ESP 协议数据结构

在隧道模式下，ESP 协议只是对整个 IP 数据包及 ESP 协议报尾部分进行加密处理，如图 14.14 所示。

图 14.14　隧道模式下 ESP 协议加密内容

隧道模式下的 ESP 协议应用如图 14.15 所示。

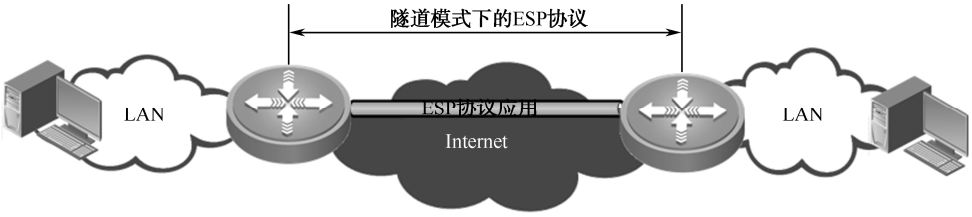


图 14.15　隧道模式下的 ESP 协议应用

3. 报文摘要 MD5

报文摘要 MD（Message Digest）是进行报文身份验证及完整性验证的简单方法。为了描述报文摘

要 MD 的工作原理，假设用户 A 将一个较长的报文 X 加密后，发送给用户 B，如图 14.16 所示。具体工作过程描述如下。

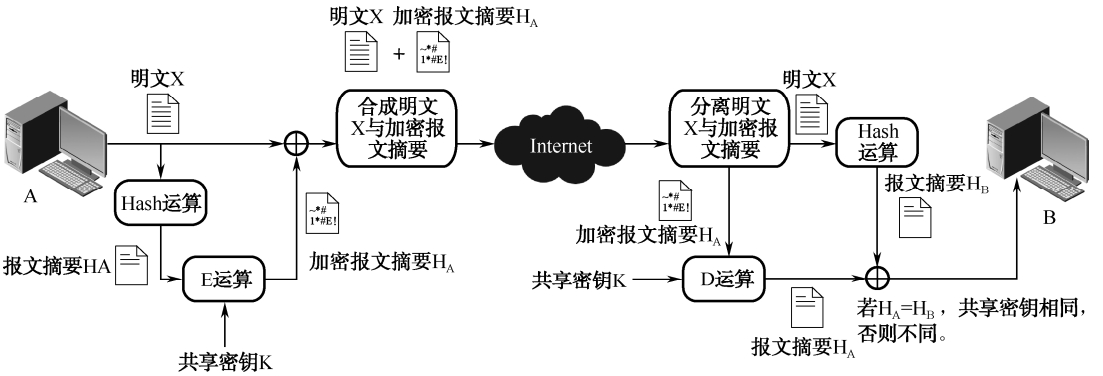


图 14.16　报文摘要 MD 身份验证与完整性验证

第 1 步：用户 A 将较长的报文 X 经过报文摘要算法运算后，得出很短的报文摘要 H_A。

第 2 步：用自己的共享密钥 K 对报文摘要 H_A 加密 E 运算，得出加密报文摘要 H_A。

第 3 步：将加密报文摘要 H_A 放置在报文 X 后面合成 一个新的报文“明文 X+加密报文摘要 H_A”，并发送给用户 B。

第 4 步：用户 B 收到合成报文“明文 X+加密报文摘要 H_A”后，将其重新分离成明文 X 及加密报文摘要 H_A。

第 5 步：用户 B 共享密钥 K 对加密报文摘要 H_A 进行解密 D 运算，得出报文摘要 H_A。

第 6 步：用户 B 同时要对分离出的报文 X 进行报文摘要运算，得出很短的报文摘要 H_B。

第 7 步：比较在用户 A 侧计算的摘要 H_A 与在用户 B 侧计算的摘要 H_B 是否一致。如果一致，说明收到的报文 X 是用户 A 的报文，且没有被修改过；否则收到的报文 X 不是用户 A 的报文或被修改过。

关于以上工作过程的几点说明如下。

（1）采用报文摘要算法是一种散列算法 Hash，它是一个单向函数。一个很长的报文 X 经过报文摘要散列算法后能够得出唯一的较短的报文摘要值 H（有时也将报文摘要值 H 称为 Hash 值）；反之，从报文摘要 H 是不能推导出报文 X 的。

（2）由于报文摘要 H 较短，对于报文摘要 H 进行加密运算，比对较长的报文 X 进行加密运算要节省很多资源，要容易得多。

（3）由于报文 X 与报文摘要 H 具有唯一性、单向性的特点，对较短的报文摘要 H 加密运算就相当于对较长的报文 X 加密运算，二者效果一样。

目前，报文摘要算法 MD5 得到广泛应用。MD5 散列算法得出的报文摘要长度为 128bit。另一种安全散列算法 SHA（Secure Hash Algorithm）与 MD5 相似，但得出的 Hash 值长度为 160bit，更加安全。

4. DH 算法

DH 算法是 Diffie-Hellman 密钥交换协议的缩写，是基于离散对数实现密钥传输的。那么 DH 是如何实现密钥传输的呢？假设用户 A 和 B 是需要获得共享密钥的两个客户。在通信前，用户 A 和 B 双方约定 2 个大整数 n 和 g，其中 1<g<n，这两个整数可以公开。

（1）用户 A 随机产生一个自然数 a，a 对外保密，然后计算 $K_a=g^a \bmod n$，将 K_a 发送给用户 B。

（2）用户 B 随机产生一个自然数 b，b 对外保密，然后计算 $K_b=g^b \bmod n$，将 K_b 发送给用户 A。

（3）用户 A 收到 B 发送的 K_b 后，计算出密钥 $K=K_b^a \bmod n=(g^b \bmod n)^a \bmod n=g^{(b*a)} \bmod n$。

（4）用户 B 收到 A 发送的 K_a 后，计算出密钥 $K=K_a^b \bmod n=（g^a \bmod n）^b \bmod n= g^{(a*b)} \bmod n$。

由于 $g^{(b*a)} \bmod n = g^{(a*b)} \bmod n$，因此可以保证双方得到的 K 是相同的，K 即是共享的密钥，如图 14.17 所示。

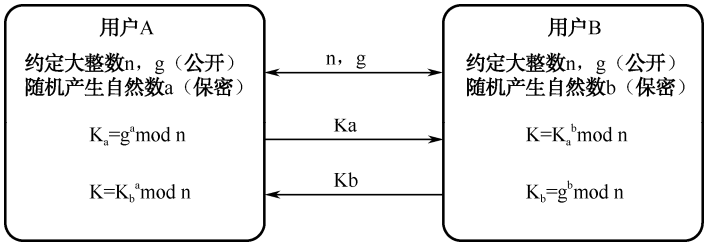


图 14.17　DH 工作原理

14.4　演示实例

1. 背景描述

假设北京的某公司在上海开了新的分公司，分公司要远程访问总公司的各种网络资源。在 Internet 上传输数据本身存在安全隐患，这家公司希望通过 IPSec VPN 技术实现数据的安全传输。由于总公司和分公司之间需要共享路由信息，所以还要使用 GRE，如图 14.18 所示。

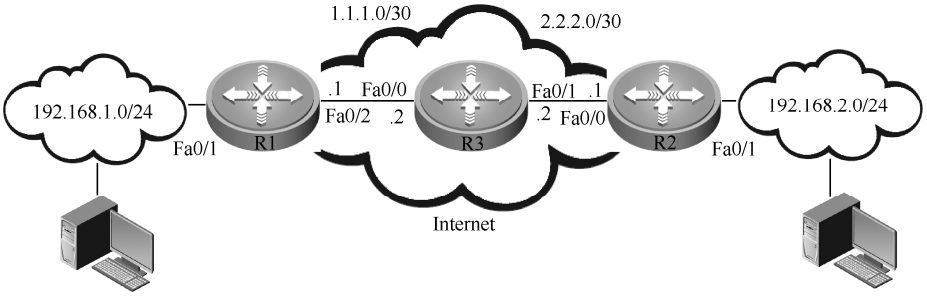


图 14.18　IPSec 应用（二）

IPSec VPN 技术通过隧道技术、加解密技术、密钥管理技术、和验证技术有效地保证了数据在 Internet 网络传输的安全性，是目前最安全、使用最广泛的 VPN 技术。GRE 支持对组播和广播数据的封装，可用于封装路由协议报文。因此我们可以通过建立 GRE over IPSec VPN 的加密隧道，实现分公司和总公司之间的路由信息共享和信息安全传输。

2. 操作步骤

（1）路由器基本配置。

第 1 步：配置 Internet 路由器 R3。

```
R3(config)#interface fastEthernet 0/0
R3(config-if)#ip address 1.1.1.2 255.255.255.252
R3(config-if)#exit
R3(config)#interface fastEthernet 0/1
R3(config-if)#ip address 2.2.2.2 255.255.255.252
R3(config-if)#exit
```

第 2 步：配置路由器 R1 与 R2 的 Internet 连通性。

```
R1(config)#interface fastEthernet 0/0
R1(config-if)#ip address 1.1.1.1 255.255.255.252
R1(config-if)#exit
R1(config)#interface fastEthernet 0/1
R1(config-if)#ip address 192.168.1.1 255.255.255.0
R1(config-if)#exit
R1(config)#ip route 0.0.0.0 0.0.0.0 1.1.1.2

R2#configure terminal
R2(config)#interface fastEthernet 0/0
R2(config-if)#ip address 2.2.2.1 255.255.255.252
R2(config-if)#exit
R2(config)#interface fastEthernet 0/1
R2(config-if)#ip address 192.168.2.1 255.255.255.0
R2(config-if)#exit
R2(config)#ip route 0.0.0.0 0.0.0.0 2.2.2.2
```

（2）GRE 隧道及 RIP 配置。

第 1 步：配置路由器 R1 的 GRE 隧道。

```
R1(config)#interface Tunnel0
R1(config-if)#ip address 10.1.1.1 255.255.255.0
R1(config-if)#tunnel source fastEthernet 0/0
R1(config-if)#tunnel destination 2.2.2.1
R1(config-if)#tunnel key 123456
R1(config-if)#exit
```

第 2 步：在路由器 R1 上启用 RIPv2 路由协议。

```
R1(config)#router rip
R1(config-router)#version 2
R1(config-router)#no auto-summary
R1(config-router)#network 10.0.0.0
R1(config-router)#network 192.168.1.0
R1(config-router)#exit
```

第 3 步：配置路由器 R2 的 GRE 隧道。

```
R2(config)#interface Tunnel0
R2(config-if)#ip address 10.1.1.2 255.255.255.0
R2(config-if)#tunnel source fastEthernet 0/0
R2(config-if)#tunnel destination 1.1.1.1
R2(config-if)#tunnel key 123456
R2(config-if)#exit
```

第 4 步：在路由器 R2 上启用 RIPv2 路由协议。

```
R2(config)#router rip
R2(config-router)#version 2
R2(config-router)#no auto-summary
```

```
R2(config-router)#network 10.0.0.0
R2(config-router)#network 192.168.2.0
R2(config-router)#exit
```

（3）IPSec 隧道配置。

第 1 步：配置路由器 R1 的 IKE 参数

配置 IKE 安全策略：

```
R1(config)#crypto isakmp policy 1
R1(isakmp-policy)#encryption 3des
R1(isakmp-policy)#authentication pre-share
R1(isakmp-policy)#hash sha
R1(isakmp-policy)#group 2
R1(isakmp-policy)#exit
```

配置 IKE 预共享密钥如下：

```
R1(config)#crypto isakmp key 0 abc123 address 2.2.2.1
```

第 2 步：配置路由器 R1 的 IPSec 参数。

配置 IPSec 转换集，使用 ESP 协议，3DES 算法和 SHA-1 散列算法，工作模式为传输模式：

```
R1(config)#crypto ipsec transform-set 3des_sha esp-3des esp-sha-hmac
R1(cfg-crypto-trans)#mode transport
R1(cfg-crypto-trans)#exit
```

配置加密访问控制列表，使用 GRE 协议（47）作为触发流量：

```
R1(config)#access-list 100 permit 47 host 1.1.1.1 host 2.2.2.1
```

配置 IPSec 访问映射条目：

```
R1(config)#crypto map to_r2 1 ipsec-isakmp
R1(config-crypto-map)#match address 100
R1(config-crypto-map)#set transform-set 3des_sha
R1(config-crypto-map)#set peer 2.2.2.1
R1(config-crypto-map)#exit
```

将访问映射条目应用于 Fa0/0 接口：

```
R1(config)#interface fastEthernet 0/0
R1(config-if)#crypto map to_r2
R1(config-if)#exit
```

第 3 步：配置路由器 R2 的 IKE 参数。

```
R2(config)#crypto isakmp policy 1
R2(isakmp-policy)#encryption 3des
R2(isakmp-policy)#authentication pre-share
R2(isakmp-policy)#hash sha
R2(isakmp-policy)#group 2
R2(isakmp-policy)#exit
R2(config)#crypto isakmp key 0 abc123 address 1.1.1.1
```

第 4 步：配置路由器 R2 的 IPSec 参数。

```
R2(config)#crypto ipsec transform-set 3des_sha esp-3des esp-sha-hmac
R2(cfg-crypto-trans)#mode transport
R2(cfg-crypto-trans)#exit
```

```
R2(config)#access-list 100 permit 47 host 2.2.2.1 host 1.1.1.1
R2(config)#crypto map to_r1 1 ipsec-isakmp
R2(config-crypto-map)#match address 100
R2(config-crypto-map)#set transform-set 3des_sha
R2(config-crypto-map)#set peer 1.1.1.1
R2(config-crypto-map)#exit
R2(config)#interface fastEthernet 0/0
R2(config-if)#crypto map to_r1
R2(config-if)#exit
```

（4）网络测试。

第 1 步：配置主机 PC1 和 PC2。

主机 PC1 的 IP 地址为 192.168.1.2，网关为 192.168.1.1。

主机 PC2 的 IP 地址为 192.168.2.2，网关为 192.168.2.1。

第 2 步：验证测试。

在主机 PC1 上 Ping PC2，可以 ping 通。

测试成功！GRE over IPSec VPN 隧道建立成功！

14.5 训练任务

某公司通过 Internet 将北京总公司与上海分公司连接，为了保证两地公司间传输数据安全，建立 IPSec 隧道。同时为保障 RIP 路由信息交换，同时启动 GRE。

要求：

1．绘制网络拓扑图；

2．路由器配置，包括基本配置、路由配置、IPSec 配置、GRE 配置等。通过查看路由表情况，验证 IPSec 及 GRE 配置是否正确。

练习题

简答题

（1）说明 IPSec 主要解决什么问题及其应用环境。

（2）描述 IPSec 的工作原理。

（3）说明 IPSec 的配置命令及步骤。

（4）比较 IPSec 与 GRE 的特性，说明它们的异同点。

参考文献

【1】张选波，吴丽征，周金玲. 设备调试与网络优化学习指南. 北京：科学出版社，2009.
【2】张选波，王东，张国清. 设备调试与网络优化实验指南. 北京：科学出版社，2009.
【3】谢希仁. 计算机网络（第 5 版）. 北京：电子工业出版社，2008.
【4】戴有炜. Windows Server 2008 R2 安装与管理. 北京：清华大学出版社，2011.
【5】RG-RSR20-04 路由器 RGOS_10.3(5b6)版本命令手册，锐捷公司，2014.
【6】RG-RSR20-04 路由器 RGOS_10.3(5b6)版本配置手册，锐捷公司，2014.
【7】RG-S3760 系列交换机 RGOS 10.3(5b1)版本命令手册，锐捷公司，2014.
【8】RG-S3760 系列交换机 RGOS 10.3(5b1)版本配置手册，锐捷公司，2014.